

10074

# The Telecommunications and Data Acquisition Progress Report 42-104

October-December 1990

E. C. Posner  
Editor

February 15, 1991



National Aeronautics and  
Space Administration

Jet Propulsion Laboratory  
California Institute of Technology  
Pasadena, California

(NASA-CP-1991-18306) THE TELECOMMUNICATIONS AND  
DATA ACQUISITION REPORT Progress Report,  
Oct. - Dec. 1990 (JPL) 273 p CSCL 17B

N91-18306  
--THRU--  
N91-18326  
Unclass  
0332658

03/32

7

# The Telecommunications and Data Acquisition Progress Report 42-104

October–December 1990

E. C. Posner  
Editor

February 15, 1991



National Aeronautics and  
Space Administration

**Jet Propulsion Laboratory**  
California Institute of Technology  
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

## Preface

This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) in planning, in supporting research and technology, in implementation, and in operations. Also included is standards activity at JPL for space data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Operations (OSO). The TDA Office also performs work funded by two other NASA program offices through and with the cooperation of the Office of Space Operations. These are the Orbital Debris Radar Program (with the Office of Space Station) and 21st Century Communication Studies (with the Office of Aeronautics and Exploration Technology).

In the search for extraterrestrial intelligence (SETI), *The TDA Progress Report* reports on implementation and operations for searching the microwave spectrum. In solar system radar, it reports on the uses of the Goldstone Solar System Radar for scientific exploration of the planets, their rings and satellites, asteroids, and comets. In radio astronomy, the areas of support include spectroscopy, very long baseline interferometry, and astrometry. These three programs are performed for NASA's Office of Space Science and Applications (OSSA), with the Office of Space Operations funding DSN operational support.

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund that involve the TDA Office are included.

This and each succeeding issue of *The TDA Progress Report* will present material in some, but not necessarily all, of the following categories:

### OSO Tasks:

- DSN Advanced Systems
  - Tracking and Ground-Based Navigation
  - Communications, Spacecraft-Ground
  - Station Control and System Technology
  - Network Data Processing and Productivity
- DSN Systems Implementation
  - Capabilities for Existing Projects
  - Capabilities for New Projects
  - New Initiatives
  - Network Upgrade and Sustaining
- DSN Operations
  - Network Operations and Operations Support
  - Mission Interface and Support
  - TDA Program Management and Analysis
- Ground Communications Implementation and Operations
- Data and Information Systems
- Flight-Ground Advanced Engineering
- Long-Range Program Planning

### OSO Cooperative Tasks:

- Orbital Debris Radar Program
- 21st Century Communication Studies

**OSSA Tasks:**

Search for Extraterrestrial Intelligence

Goldstone Solar System Radar

Radio Astronomy

**Discretionary Funded Tasks**

# Contents

## OSO TASKS DSN Advanced Systems TRACKING AND GROUND-BASED NAVIGATION

<b>Demonstration of Precise Estimation of Polar Motion Parameters With the Global Positioning System: Initial Results</b> .....	1
S. M. Lichten	51
NASA Code 310-10-61-84-04	

## COMMUNICATIONS, SPACECRAFT-GROUND

<b>Analysis of a Thick Dichroic Plate With Rectangular Holes at Arbitrary Angles of Incidence</b> .....	9
J. C. Chen	52
NASA Code 310-20-64-89-01	
<b>An Analysis of the Least-Squares Problem for the DSN Systematic Pointing Error Model</b> .....	17
L. S. Alvarez	53
NASA Code 310-20-65-63-00	
<b>Predictive Control and Estimation Algorithms for the NASA/JPL 70-Meter Antennas</b> .....	30
W. Gawronski	54
NASA Code 310-20-65-63-00	
<b>Initial Studies of Array Feeds for the 70-Meter Antenna at 32 GHz</b> .....	50
P. W. Cramer	55
NASA Code 310-20-65-89-01	
<b>Steady-State Phase Error for a Phase-Locked Loop Subjected to Periodic Doppler Inputs</b> .....	68
C.-C. Chen and M. Z. Win	56
NASA Code 310-20-67-59-00	
<b>Maximum-Entropy Probability Distributions Under <math>L_p</math>-Norm Constraints</b> .....	74
S. Dolinar	57
NASA Code 310-30-71-83-02	
<b>Compressed/Reconstructed Test Images for CRAFT/Cassini</b> .....	88
S. Dolinar, K.-M. Cheung, I. Onyszchuk, F. Pollara, and S. Arnold	58
NASA Code 310-30-71-83-02	
<b>Emerging Standards for Still Image Compression: A Software Implementation and Simulation Study</b> .....	98
F. Pollara and S. Arnold	59
NASA Code 310-30-71-83-02	
<b>Algebraic Geometric Codes</b> .....	103
M. Shahshahani	510
NASA Code 310-30-71-83-02	

## STATION CONTROL AND SYSTEM TECHNOLOGY

<b>Overview of Arraying Techniques in the Deep Space Network</b> .....	109
A. Mileant and S. Hinedi	511
NASA Code 310-30-70-56-01	

<b>The Advanced Receiver II: Telemetry Test Results in CTA 21</b> .....	140
S. Hinedi, R. Bevan, and M. Marina	
NASA Code 310-30-71-84-02	

## DSN Systems Implementation SUSTAINING ENGINEERING

<b>On Estimating the Phase of a Periodic Waveform in Additive Gaussian Noise— Part III</b> .....	157
L. L. Rauch	
NASA Code 314-40-31-10-83	

## CAPABILITIES FOR EXISTING PROJECTS

<b>Digital Test Signal Generation: An Accurate SNR Calibration Approach for the DSN</b> .....	161
B. O. Gutierrez-Luaces	
NASA Code 314-40-41-11-06	
<b>A Comparison of Manchester Symbol Tracking Loops for Block V Applications</b> .....	175
J. K. Holmes	
NASA Code 314-40-42-01-39	
<b>Open-Loop Frequency Acquisition for Suppressed-Carrier Biphase Signals Using One-Pole Arm Filters</b> .....	192
B. Shah and J. K. Holmes	
NASA Code 314-40-42-01-39	
<b>Frame Synchronization for the Galileo Code</b> .....	211
S. Arnold and L. Swanson	
NASA Code 314-30-51-01-15	
<b>Planetary Ephemerides Approximation for Radar Astronomy</b> .....	219
R. Sadr and M. Shahshahani	
NASA Code 314-30-63-31-02	

## DSN Operations Flight–Ground Advanced Engineering

<b>New Space Research Frequency Band Proposals in the 20- to 40.5-GHz Range</b> .....	235
D. F. Bishop	
NASA Code BG315-20-50-00-01	
<b>Design Concepts and Performance of NASA X-Band (7162-MHz/8415-MHz) Transponder for Deep-Space Spacecraft Applications</b> .....	247
N. R. Mysoor, J. D. Perret, and A. W. Kermod	
NASA Code 314-30-55-30-01	
<b>Author Index, 1990</b> .....	257

S1-04  
P-8  
N91-18309 U50

# Demonstration of Precise Estimation of Polar Motion Parameters With the Global Positioning System: Initial Results

S. M. Lichten

Tracking Systems and Applications Section

*Data from the Global Positioning System (GPS) have been used to determine precise polar motion estimates. Conservatively calculated formal errors of the GPS least-squares solution are approximately 10 cm. The GPS estimates agree with independently determined polar motion values from very long baseline interferometry (VLBI) at the 5-cm level. The data were obtained from a partial constellation of GPS satellites and from a sparse worldwide distribution of ground stations. The accuracy of the GPS estimates should continue to improve as more satellites and ground receivers become operational, and eventually a near-real-time GPS capability should be available. Because the GPS data are obtained and processed independently from the large radio antennas at the Deep Space Network (DSN), GPS estimation could provide very precise measurements of Earth orientation for calibration of deep space tracking data and could significantly relieve the ever-growing burden on the DSN radio telescopes to provide Earth platform calibrations.*

## I. Introduction

Precise navigation and tracking for high Earth orbiter and interplanetary missions require knowledge of certain Earth platform parameters, including Earth orientation. Earth orientation [1] parameters include the angle (and rate) of rotation of the Earth relative to a reference position, and the position of the solid crust and mantle relative to the axis of rotation. These parameters can vary unpredictably on a daily (or more frequent) basis due to interactions between the rotating solid Earth, its oceans, and atmosphere. By incorporating measurements made over

several days and by using a Kalman filter, it is possible to smooth and predict the longer-term components of Earth orientation. Presently, Earth orientation calibrations for deep-space tracking are made twice a week by using combinations of 70-m and 34-m antennas for very long baseline interferometry (VLBI) observations of quasar radio sources. Up to 10 hr per week can be required for this task for support of Magellan; similar demands are expected for the Galileo mission. Spacecraft telemetry communication is generally not possible when Deep Space Network (DSN) antennas are used for these Earth orientation observation sessions. Although the present-day VLBI technique ap-

appears sufficient to meet the 30-cm Magellan/Galileo-Earth orientation requirement (needed within about one day after the data are taken), the technique leaves less time available for direct spacecraft tracking. If interplanetary missions to the Moon, Mars, and other destinations occur later in the 1990s, as expected, DSN radio antennas will become increasingly busy, and antenna time will be at a premium.

A relatively new technique for monitoring Earth orientation incorporates data from the United States Global Positioning System (GPS), which presently includes about 15 navigation satellites, to be expanded to 18 by 1992, and further to 24 satellites by the mid-1990s. As described in [2], the GPS data will be combined with VLBI measurements. In this combined system, VLBI observations can be made with greatly reduced frequency from the present. The result of the combined GPS/VLBI system is expected to be an Earth orientation monitoring technique that can be made more accurate than the original VLBI system alone, but requires significantly less DSN radio antenna time. In addition to enhancing productivity of the DSN by enabling more time for direct spacecraft tracking and telemetry communication, the improved accuracy of the new system could help the DSN better support a sub-nanoradian deep-space tracking capability.

The DSN has installed advanced GPS terminals at each deep space tracking site for ionospheric calibrations.<sup>1</sup> Since the GPS transmits at two L-band frequencies, the lowest order ionospheric path delays can be determined straightforwardly at each site by using GPS data. These DSN GPS receivers will soon become the reference sites for a worldwide, high-precision GPS tracking network. The TOPEX/POSEIDON oceanography satellite, carrying its own GPS receiver, will be tracked with GPS differential techniques by this worldwide network. For several years, data from other, less precise worldwide GPS stations have been collected and processed. In order to test the capability of such networks to provide measurements of Earth orientation, data from one of the earliest worldwide GPS experiments in 1988 were processed to estimate X and Y polar motion parameters. The results of this test are reported in this article.

## II. Experimental Data

The GPS experiment utilized for this study is described in [3] and [4]. To estimate Earth orientation parameters,

<sup>1</sup> C. J. Vegos, *DSCC Media Calibration Subsystem (DMD), Functional Design Review (Level D)*, JPL 834-30, vol. 1 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 1, 1987.

a three-day arc (January 19–21, 1988) was selected from the experiment, which spanned about three weeks. This particular three-day arc actually did not contain an optimal set of measurements, since one of the four fiducial (fixed reference) sites did not collect data on the first day. However, these data were the first to be processed and conveniently available for use. The four fiducial sites were all at VLBI observatories where precise coordinates were available: Hatcreek, California; Fort Davis, Texas; Haystack, Massachusetts; and Onsala, Sweden. The coordinates of the other sites were estimated from the GPS data: Wettzell, Germany; Canberra, Australia; Black Birch, New Zealand; Limón, Costa Rica; Liberia, Costa Rica; Cocos Island (off South America in the Pacific Ocean); Mojave, California; and Owens Valley, California. The data from these 12 stations represent less than half of the data actually collected over the three-day arc. Additional stations were left out in order to reduce the computational load, since all the GPS data were reduced on microcomputers. Figure 1 shows the location of the 12 sites used and the sparse worldwide coverage that was available. The satellite constellation in 1988 included only seven satellites, which provided only partial geometrical coverage.

## III. Approach and Results

GPS-based positioning techniques generally involve a least-squares estimation of various parameters from the GPS data. For high-precision applications, such as the determination of Earth orientation, the parameters include the GPS orbits themselves. Other estimated parameters include nonfiducial station coordinates; relative clock offsets of GPS receivers and transmitters modeled as a white-noise process from measurement to measurement; zenith troposphere delays at each site, modeled as a random-walk stochastic process, with noise level of  $(1.2 \text{ cm})^2/\text{hr}$ ; the Earth orientation polar motion parameters; and GPS carrier phase bias parameters. All parameters were estimated simultaneously.

The a priori uncertainties of all the parameters, other than the Earth orientation parameters, were very large, effectively infinite. The a priori uncertainty of the polar motion parameters was initially 50 cm. Since the final formal sigmas for these parameters ended up at the level of  $\sim 10$  cm, the run was repeated with 200-cm a priori sigmas, but the solutions changed by less than 0.2 cm, confirming that the a priori sigmas for the polar motion parameters were not constraining the GPS solutions.

The solutions were determined with a square-root information filter. The method is described in [5–7]. The

data were compressed to 6-min intervals. Data weights used were 1 cm for the GPS carrier phase and 200 cm for the GPS pseudorange. As described in [8], the GPS carrier phase provides a very precise measure of range change but is ambiguous by an integer multiple of wavelengths. By processing together the phase and range data and by using an innovative technique developed for resolving the carrier phase ambiguities [8], most of the carrier phase biases in North and South America were resolved. Because of correlations between the phase biases and the clock parameters, resolving the ambiguities is possible only between pairs of stations and satellites. The greater the distance between the two stations, the more difficult it is to resolve ambiguities involving that pair. In principle, when carrier phase biases are resolved, solutions should improve significantly, since the ambiguous phase data (measuring *range change*) has been effectively converted to a very precise (subcentimeter) *range* measurement. One of the goals of this study was to examine the effect of GPS phase ambiguity techniques on the estimation of global Earth orientation parameters. Although ambiguity resolution has been known to dramatically improve the accuracy of baseline estimates [8], the effect on global Earth platform parameters has not been studied prior to the analysis presented in this article.

The results appear in Table 1. The differences between the GPS and VLBI (International Radio Interferometric Surveying, IRIS) estimates for X and Y polar motion are listed along with the formal errors from the GPS solution. Most (~70 percent) of the GPS formal errors, which are simply the parameter standard deviations from the covariance matrix, are from the computed error (data noise, geometry, and satellite visibility), with the remainder due to systematic error from an assumed 4-cm error for each fiducial station coordinate. The systematic fiducial coordinate errors were calculated from a consider analysis [9] and are probably conservative, since GPS and VLBI comparisons, at least in North America, have shown that the GPS fiducial coordinate errors are probably at the level of ~2 cm or better per coordinate [5]. Note that after ambiguity resolution, the GPS and VLBI Earth orientation estimates agree to 5 cm or better in each component.

#### IV. Discussion

The reader is cautioned that the comparison of polar motion parameters from GPS and VLBI techniques should be regarded as preliminary, since only a three-day GPS solution was used. However, previous results reported elsewhere [10] with somewhat cruder GPS orbit determination strategies have shown consistency between VLBI and GPS at the 10–20-cm level during a two-year period. The other

results [10] utilized a weaker GPS orbit solution, with ~5-m accuracy. The orbits reported here have been improved to the level of 60–100 cm [7] through careful modeling and orbit estimation techniques, so Earth orientation accuracy of ~5 cm is indeed consistent with results reported elsewhere when scaled for the different orbit accuracies (assuming that the errors are proportional).

The formal errors shown in Table 1 are large enough (~10 cm) so that there is no statistically significant difference between the GPS and VLBI polar motion estimates. In fact, the accuracy of the VLBI estimates themselves is presently believed to be ~3 cm, which is not much different from the 5-cm differences observed. The GPS-VLBI (IRIS) polar motion differences are actually a measure of the offset between the reference frame defined by the GPS fiducial site coordinates (which were held fixed in the solution) and the IRIS Earth orientation time series. Since both the fiducial site coordinates and the IRIS Earth orientation values were determined with VLBI, it might be expected a priori that the offset should be zero. However the fiducial site coordinates were obtained from a global Crustal Dynamics Project (CDP) VLBI solution (solution set GLB223 from Goddard Space Flight Center), which is not entirely consistent with the IRIS Earth orientation time series: an intercomparison in [11] showed that 1–3-cm discrepancies between independent VLBI solutions (CDP and IRIS) for polar motion are observed. Although the present GPS estimates are probably not sufficiently precise to detect this discrepancy, it is clear that a unified reference frame must be defined so that GPS Earth orientation estimates that incorporate VLBI fiducial coordinates can be related to VLBI estimates of Earth orientation.

Table 1 also shows a dramatic improvement in the accuracy of the GPS polar motion estimates with ambiguity resolution. This improvement is confirmed both in the formal calculated errors and in the comparison with independent VLBI estimates.<sup>2</sup> The result is intriguing since polar motion is essentially a global quantity, while GPS phase ambiguity resolution is most effective for shorter baselines. In fact, in this experiment, no ambiguities were resolved between the continents of North America and Europe (more than 5000 km apart), although most ambiguities of a few thousand kilometers or less over baselines in North and South America were resolved. Thus, ambiguity resolution within local continental networks was still able to significantly improve estimates of global quantities. This process can be understood in the context of

<sup>2</sup> Although the Y polar motion GPS-VLBI difference appears to be slightly worse, the difference is, in fact, statistically insignificant.

improving the GPS orbits through ambiguity resolution: even relatively short distance ambiguity resolution can improve the accuracy of the orbits. These improved orbits, in turn, result in better estimates for other, global parameter estimates. There may be implications for strategies to determine global Earth platform parameters with GPS if results, such as those in Table 1, are confirmed in future studies. By incorporating GPS data from a few stations relatively close (approximately hundreds to a few thousand kilometers) to the DSN fiducial GPS sites, GPS estimates of Earth orientation may be considerably enhanced. Data from these additional sites are fairly easy to acquire. For example, in California, there are dozens of GPS sites from which data are routinely collected by several U.S. Government agencies. Similar data are now or soon will be available from Europe and the South Pacific as GPS ground networks for various National Aeronautics and Space Administration (NASA) and international geodetic programs densify and GPS ground stations for other missions (such as TOPEX/POSEIDON and the Earth Observing System, EOS) became operational.

Other facets of GPS technology applicable to Earth orientation include the potential for a near-real-time capability and support for very high-precision ( $\sim 1$  nanoradian) VLBI deep space tracking/navigation. An advanced version of the DSN GPS Rogue receiver is in development that will automatically perform numerous data reduction steps in the receiver while the data are collected in the field.<sup>3</sup> This would enable very fast data turnaround and the potential for routine delivery of Earth platform parameters the same (or next) day. Certain data types used by the DSN, such as differential Doppler, require accurate Earth orientation calibrations to achieve their full navigation potential.<sup>4,5</sup> A GPS network in California is presently being tested for rapid production of GPS orbits and other related parameters. A recent test using receivers in California resulted in GPS orbits estimated in less than 24 hours after the data were acquired in the field. The ultra-precise (1-nanoradian or better) VLBI deep space tracking system currently being studied at JPL [12] under DSN Advanced Systems for future planetary missions

incorporates the estimation of Earth platform parameters in the course of multiple VLBI observations of quasars and spacecraft. Centimeter-level Earth orientation knowledge from GPS could constrain the a priori uncertainties in the Earth platform parameters to be further refined and improved from the VLBI data. Such a priori knowledge would enhance the performance of the system since it would require less data (and therefore fewer DSN resources) to achieve the desired accuracy.

As discussed in [2], GPS data can be quite sensitive to the rate of change of UT1-UTC, also known as the length of day. This information could be combined with relatively infrequent VLBI measurements of UT1-UTC to provide the DSN with a continuous and accurate record of UT1-UTC, a quantity that is needed along with polar motion for deep space navigation. However, the sparse ground coverage and relatively small number of satellites available during the 1988 GPS experiment hindered the determination of the length of day (the UT1-UTC rate), even with the three-day arc used. The three-day arc GPS estimate of the UT1-UTC rate had formal uncertainty of more than 5 cm/day. In future experiments with more satellites and ground stations operating, it should be possible to demonstrate a UT1-UTC rate accuracy of better than 1 cm/day along with accuracy for pole position improving to the level of 1-2 cm. One such experiment is expected to take place in early 1991.

## V. Summary and Conclusions

An initial GPS solution for X and Y polar motion agrees with VLBI to better than 5 cm in each component. The GPS experiment was successful beyond expectations, with GPS-VLBI agreement better than the  $\sim 10$ -cm-level predicted from the GPS solution covariance. Even more significant, perhaps, is the fact that the GPS solution was obtained with a somewhat sparse ground network tracking only seven GPS satellites, about one-third of the full GPS constellation, which is expected to be complete by 1992.

The GPS polar motion estimates improved significantly after GPS phase ambiguity resolution. This is the first demonstration that global Earth parameters estimated from GPS data improve when ambiguity resolution is performed, even though GPS ambiguity resolution is generally carried out over baselines much shorter than the radius of the Earth.

Further improvement in GPS polar motion estimates to the few-centimeter level is eventually expected as advanced

<sup>3</sup> W. B. Melbourne, J. B. Thomas, C. L. Thornton, L. E. Young, and T. P. Yunck, *Turbo Rogue: The Next Generation GPS Geodetic Receiver* (internal document), Jet Propulsion Laboratory, Pasadena, California, November 9, 1989.

<sup>4</sup> S. W. Thurman, "Deep-Space Navigation Performance of X-Band (8.4 GHz) Differenced Doppler Data," JPL Interoffice Memorandum 314-486 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 24, 1990.

<sup>5</sup> S. W. Thurman, "DSN Baseline Coordinate and Station Location Errors Induced by Earth Orientation Errors," JPL Interoffice Memorandum 314-488 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 25, 1990.

DSN-type GPS receivers are installed worldwide, the GPS constellation begins to fill out, and the receivers are spread out more evenly around the globe. Additional GPS experiments are planned in the coming years to monitor this improvement, confirm the results reported in this article, and to demonstrate estimation of polar motion rates and the UT1-UTC rate with GPS.

The potential benefits to the DSN of using GPS for Earth platform parameter estimation are becoming clear with demonstrations of the inherently high precision and high data density that are available with GPS. The GPS

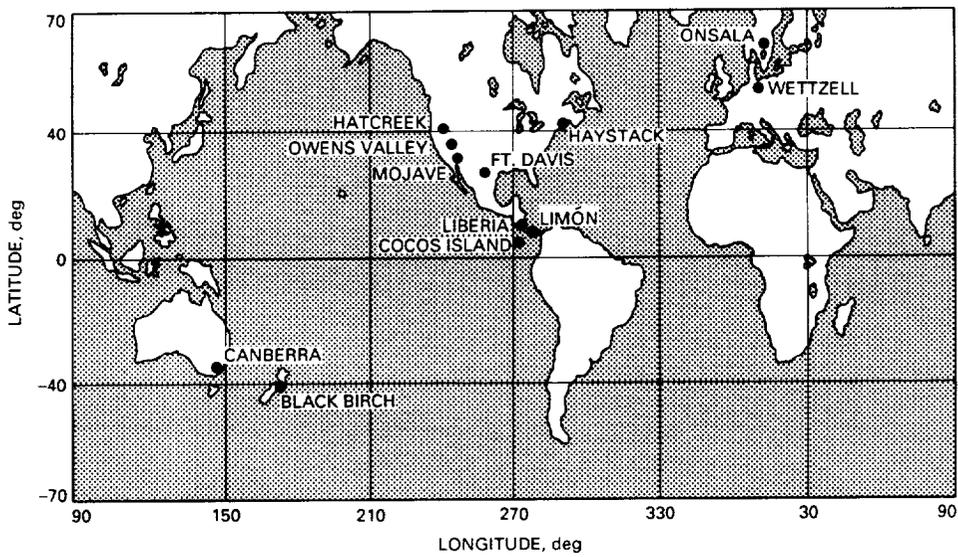
measurements can be combined with DSN VLBI observations to produce a time history of Earth orientation and rotation with high accuracy and high temporal resolution, requiring only a fraction of the antenna time that must presently be allocated to measure these quantities with VLBI alone. In addition to conserving DSN resources, GPS techniques could eventually result in a near-real-time estimation capability of Earth orientation for the DSN. The centimeter-level Earth orientation accuracy expected from the GPS data could also support and enhance an ultra-precise (subnanoradian) VLBI-based tracking and navigation system for the DSN.

## References

- [1] K. Lambeck, *The Earth's Variable Rotation: Geophysical Causes and Consequences*, Cambridge, England: Cambridge University Press, 1980.
- [2] A. P. Freedman, "Determination of Earth Orientation Using the Global Positioning System," *TDA Progress Report 42-99*, vol. July-September 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 1-11, November 15, 1989.
- [3] J. N. Kellogg and T. H. Dixon, "Central and South America GPS Geodesy—CASA UNO," *Geophysical Research Letters*, vol. 17, no. 3, pp. 195-198, March 1990.
- [4] R. E. Neilan, T. H. Dixon, T. K. Meehan, W. G. Melbourne, J. A. Scheid, J. N. Kellogg, and J. L. Stowell, "Operational Aspects of CASA UNO'88—The First Large Scale International GPS Geodetic Network," *IEEE Transactions on Instrumentation and Measurement*, vol. 38, pp. 648-651, April 1989.
- [5] S. M. Lichten and W. I. Bertiger, "Demonstration of Sub-Meter GPS Orbit Determination and 1.5 Parts in  $10^8$  Three-Dimensional Baseline Accuracy," *Bulletin Geodesique*, vol. 63, no. 5, pp. 167-189, 1989.
- [6] S. M. Lichten, "Estimation and Filtering for High-Precision GPS Applications," *Manuscripta Geodaetica*, vol. 15, pp. 159-176, 1990.
- [7] S. M. Lichten, "Towards GPS Orbit Accuracy of Tens of Centimeters," *Geophysical Research Letters*, vol. 17, pp. 215-218, March 1990.
- [8] G. Blewitt, "Carrier Phase Ambiguity Resolution for the Global Positioning System Applied to Geodetic Baselines up to 2000 km," *Journal of Geophysical Research*, vol. 94, pp. 10187-10203, August 10, 1989.
- [9] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, New York: Academic Press, 1977.
- [10] E. R. Swift, "Earth Orientation from GPS," presented at the 5th Annual GPS Workshop, Jet Propulsion Laboratory, Pasadena, California, April 2, 1990.
- [11] T. Yoshino, Y. Takahashi, N. Kawaguchi, K. Heki, K. Yokoyama, and S. Manabe, "Intercomparison of the Earth Rotation Parameters by Two Independent VLBI Networks," *Astronomy and Astrophysics*, vol. 224, pp. 316-320, 1989.
- [12] R. N. Treuhaft and S. T. Lowe, "A Demonstration of Nanoradian VLBI Tracking for Deep Space Navigation," paper AIAA 90-2939, AAS/AIAA Astrodynamics Conference, Portland, Oregon, pp. 587-589, August 1990.

**Table 1. Difference between GPS and VLBI Earth orientation estimates from  
January 19–21, 1988**

Parameter	Initial GPS solution (no ambiguities resolved)	Final GPS solution (with ambiguity resolution)
X pole ( $\sigma$ formal error)	+16.3 cm ( $\pm 15$ )	-3.6 cm ( $\pm 8$ )
Y pole ( $\sigma$ formal error)	-2.0 cm ( $\pm 19$ )	-4.4 cm ( $\pm 12$ )



**Fig. 1. Sites used in the 1988 worldwide GPS experiment for estimation of polar motion parameters.**

32-32  
P-8  
N91-18310  
J J 5

# Analysis of a Thick Dichroic Plate with Rectangular Holes at Arbitrary Angles of Incidence

J. C. Chen

Ground Antennas and Facilities Engineering Section

*A thick dichroic plate that is transparent to one frequency band while at the same time reflective to other frequencies is needed for high-power transmission in the Deep Space Network. Software based on the method of moments was developed to design a thick dichroic plate with rectangular holes. A test dichroic plate was fabricated, and an experiment was performed to verify the computer code. Good agreement was found between theory and experiment.*

## I. Introduction

The Deep Space Network has a need for dichroic plates having an insertion loss as low as 0.04 dB at the pass-band and the ability to handle up to 500 kW of power [1]. Because a thin dichroic plate may not be mechanically suitable for these requirements, a thick dichroic plate design is considered. Also, using rectangular rather than circular holes provides the designer with an extra degree of freedom. Therefore a study of a thick dichroic plate with rectangular holes was made, and a computer code was developed.

The relationship between the rectangular holes, the array lattice of an infinite dichroic plate, and the incident wave is shown in Fig. 1. The design variables of the software are hole size, lattice size and shape, thickness of the plate, dielectric constant in the hole, and angle of incidence of the linearly polarized wave. The program calculates the amplitude and phase of both the transmission and reflection coefficients of the two orthogonal linear pol-

arizations (TE and TM). The program was run on JPL's Cray X-MP/18 supercomputer.

## II. Analysis

The analysis of a thick dichroic plate with rectangular holes is carried out in a series of steps. First, a model of a half-space infinite array is constructed. A complete set of basis functions with unknown coefficients is developed for the waveguide region (waveguide modes) and for the free-space region (Floquet modes) in order to represent the electromagnetic fields [2]. Next, the boundary conditions are applied at the interface between these two regions. The method of moments is used to compute the unknown mode coefficients [3,4]. The scattering matrix of the half-space infinite array is then calculated. The reference plane of the scattering matrix is moved half a plate thickness in the negative  $z$  direction. Finally, a dichroic plate of finite thickness is synthesized by positioning two plates of half thickness back to back. The total scattering matrix is

obtained by cascading the scattering matrices of the two half-space infinite arrays.

### A. Floquet Modes and Waveguide Modes

The analysis starts with an infinite array in half space, Fig. 2(a), consisting of a free-space region ( $z > 0$ ) and a waveguide region ( $z < 0$ ). The electromagnetic fields in each region are represented by a set of orthonormal basis functions, Floquet and waveguide modes respectively, which satisfy Maxwell's equations. The Floquet modes are given by

$$\overline{\Psi}_{mn}^{TE}(x, y) = \frac{1}{\sqrt{D_x D_y \sin \Omega}} \frac{\beta_n \hat{x} - \alpha_m \hat{y}}{\sqrt{\alpha_m^2 + \beta_n^2}} e^{j(\alpha_m x + \beta_n y)} \quad (1)$$

for Floquet  $TE_{mn}$  modes, and

$$\overline{\Psi}_{mn}^{TM}(x, y) = \frac{1}{\sqrt{D_x D_y \sin \Omega}} \frac{\alpha_m \hat{x} + \beta_n \hat{y}}{\sqrt{\alpha_m^2 + \beta_n^2}} e^{j(\alpha_m x + \beta_n y)} \quad (2)$$

for Floquet  $TM_{mn}$  modes, with

$$\alpha_m = \frac{2\pi m}{D_x} - k_0 \sin \theta \cos \phi \quad (3)$$

and

$$\beta_n = \frac{2\pi n}{D_y \sin \Omega} - \frac{2\pi m}{D_x} \cot \Omega - k_0 \sin \theta \sin \phi \quad (4)$$

where  $k_0$  is the wave number in the free space.

The waveguide modes are given by

$$\begin{aligned} \overline{\Phi}_{mn}^{TE}(x, y) &= \frac{1}{\sqrt{\left(\frac{m\pi}{H_x}\right)^2 + \left(\frac{n\pi}{H_y}\right)^2}} \sqrt{\frac{\epsilon_m \epsilon_n}{H_x H_y}} \\ &\times \left[ \frac{n\pi}{H_y} \cos\left(\frac{m\pi x}{H_x} + \frac{m\pi}{2}\right) \sin\left(\frac{n\pi y}{H_y} + \frac{n\pi}{2}\right) \hat{x} \right. \\ &\quad \left. - \frac{m\pi}{H_x} \sin\left(\frac{m\pi x}{H_x} + \frac{m\pi}{2}\right) \cos\left(\frac{n\pi y}{H_y} + \frac{n\pi}{2}\right) \hat{y} \right] \quad (5) \end{aligned}$$

for rectangular waveguide  $TE_{mn}$  modes, and

$$\begin{aligned} \overline{\Phi}_{mn}^{TM}(x, y) &= \frac{1}{\sqrt{\left(\frac{m\pi}{H_x}\right)^2 + \left(\frac{n\pi}{H_y}\right)^2}} \sqrt{\frac{\epsilon_m \epsilon_n}{H_x H_y}} \\ &\times \left[ \frac{m\pi}{H_x} \cos\left(\frac{m\pi x}{H_x} + \frac{m\pi}{2}\right) \sin\left(\frac{n\pi y}{H_y} + \frac{n\pi}{2}\right) \hat{x} \right. \\ &\quad \left. + \frac{n\pi}{H_y} \sin\left(\frac{m\pi x}{H_x} + \frac{m\pi}{2}\right) \cos\left(\frac{n\pi y}{H_y} + \frac{n\pi}{2}\right) \hat{y} \right] \quad (6) \end{aligned}$$

for rectangular waveguide  $TM_{mn}$  modes, where  $\epsilon_m$  is the Neumann factor and equals 1 for  $m = 0$  and 2 for  $m > 0$ .

The periodicity of the infinite array simplifies the analysis to a study of a single equivalent element. If the problem did not have this periodicity, the mutual coupling would have to be evaluated element by element, increasing the complexity of the computation.

### B. Boundary Conditions and Method of Moments

The electromagnetic field in the waveguide is expressed as a sum of incident and reflected waveguide modes, while in the free-space region it is expressed as a sum of incident and reflected Floquet modes. Boundary conditions are applied at the interface between two regions, i.e., the transverse electric and magnetic field must be continuous across the junction at  $z = 0$ . This leads to an integral equation for the unknown transverse electric field at the boundary. The infinite-array scattering problem then becomes similar to a two-region waveguide problem.

The method of moments is used to transform the integral equation into a matrix equation suitable for evaluation on a digital computer. Solving the set of linear equations gives the unknown mode coefficients in both regions. The time required for filling the matrix depends on the number of Floquet and waveguide modes used. The number of waveguide modes and Floquet modes used in the program can be increased by the user to ensure convergence of the solution [5,6].

### C. Scattering Matrix and Reference Plane

The characteristics of the infinite array referenced to  $z = 0$  are represented by a scattering matrix  $\mathbf{S}$  which contains the transmission and reflection information for the free-space/waveguide junction.

$$\mathbf{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \quad (7)$$

$$S''_{21} = S'_{12} \quad (16)$$

$$S''_{22} = S'_{11} \quad (17)$$

where  $S_{11}$ ,  $S_{12}$ ,  $S_{21}$ , and  $S_{22}$  are matrices with 1 representing the free-space region and 2 the waveguide region. The size of matrix  $S_{11}$  is 2 by 2,  $S_{22}$  is  $n$  by  $n$ ,  $S_{12}$  is 2 by  $n$ ,  $S_{21}$  is  $n$  by 2, and  $n$  is the number of waveguide modes used. For an arbitrary set of incident waveguide modes contained in vector  $a_1$  and incident  $TE_{00}$  and  $TM_{00}$  Floquet modes contained in vector  $a_2$ , the reflected mode vectors  $b_1$  and  $b_2$  are determined by the following set of equations.

$$b_1 = S_{11}a_1 + S_{12}a_2 \quad (8)$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \quad (9)$$

Moving the reference plane in the waveguide region from  $z = 0$  to  $z = -t/2$ , Fig. 2(b), where  $t$  is the thickness of the plate, the elements of the new scattering matrix  $\mathbf{S}'$  become

$$S'_{11}(u, v) = S_{11}(u, v) \quad (10)$$

$$S'_{12}(u, v) = S_{12}(u, v) \exp\left(-j\gamma_v \frac{t}{2}\right) \quad (11)$$

$$S'_{21}(u, v) = S_{21}(u, v) \exp\left(-j\gamma_u \frac{t}{2}\right) \quad (12)$$

$$S'_{22}(u, v) = S_{22}(u, v) \exp\left[\left(-j\gamma_u \frac{t}{2}\right)\left(-j\gamma_v \frac{t}{2}\right)\right] \quad (13)$$

where  $\gamma_u$  and  $\gamma_v$  are the propagation constants of modes  $u$  and  $v$ .

#### D. Cascading and Finite Thickness

Scattering by a dichroic plate with finite thickness can be analyzed by considering two infinite-array problems. The space is divided into four regions: a free-space region (region I), two waveguide regions (regions II and III), and another free-space region (region IV), as shown in Fig. 3. The scattering matrix with reference to  $z = -t/2$  for regions I and II is  $\mathbf{S}'$ , and the scattering matrix with reference to  $z = -t/2$  for regions III and IV is  $\mathbf{S}''$ , which is the transpose matrix of  $\mathbf{S}'$ .

$$S''_{11} = S'_{22} \quad (14)$$

$$S''_{12} = S'_{21} \quad (15)$$

The scattering matrix  $\mathbf{S}^T$  for the finite-thickness plate is determined by cascading these two matrices.

$$S_{11}^T = S'_{12}(I - S''_{11}S'_{22})^{-1}S''_{11}S'_{21} + S'_{11} \quad (18)$$

$$S_{12}^T = S'_{12}(I - S''_{11}S'_{22})^{-1}S''_{12} \quad (19)$$

$$S_{21}^T = S''_{21}(I - S'_{22}S''_{11})^{-1}S'_{21} \quad (20)$$

$$S_{22}^T = S''_{21}(I - S'_{22}S''_{11})^{-1}S'_{22}S''_{12} + S''_{22} \quad (21)$$

where  $I$  is a unitary matrix.

### III. Verification of the Computer Code

The dichroic plate computer program was first checked against the available calculations for a thick dichroic plate with square holes at normal incidence [7]. The results showed good agreement. Next, a 13.25-by-15.5-in. metallic dichroic plate (Fig. 4) was fabricated to verify the computer code. The dimensions of the plate are  $H_x = 0.771$  in.,  $H_y = 0.757$  in.,  $D_x = 0.940$  in.,  $D_y = 0.940$  in.,  $\Omega = 60.0$  deg, and  $t = 1.411$  in., with a tolerance of 0.001 in. Figure 5 shows the convergence of the solution at 8.3 GHz for TM polarization with respect to the number of waveguide modes used. Satisfactory convergence is achieved by using 40 waveguide modes. The amplitude and phase of the transmission and reflection coefficients from 8 to 9 GHz were computed with 40 waveguide modes for TE and TM polarizations (Figs. 6–9). Since the plate was built merely to verify the software, it was not optimized for any specific passband.

The configuration of the experiment is shown in Fig. 10. The reflection coefficients were measured from 8 to 9 GHz for  $\theta = 30$  deg and  $\phi = 90$  deg. The feed used was a 22-dB corrugated horn with a far-field phase center 3.125 in. inside the horn aperture. A lens was added to the horn aperture to bring the far-field phase center to 2.0 in. outside the horn aperture. Since the new phase center was closer to the dichroic plate, a smaller plate with less than -25-dB edge taper was used.

The experimental results and the calculation using 40 waveguide modes are shown in Figs. 11 and 12 for TE

and TM linear polarization, respectively. The curves are in excellent agreement except for frequencies higher than 8.6 GHz for TM polarization. This area of discrepancy can be explained by considering the effects of non-plane-wave incidence.

The analysis of the thick dichroic plate assumes that a linearly polarized plane wave is incident on an infinite dichroic plate at 30 deg. In the experiment, a corrugated horn was used. The field radiated contains plane wave components that are incident at angles other than 30 deg. This can result in grating-lobe problems at high

frequencies, as can be seen in Fig. 12 for frequencies above 8.6 GHz.

#### IV. Conclusion

An analysis of a thick dichroic plate with rectangular holes has been presented. The good agreement between calculation and experimental data demonstrates the accuracy of the software. Further studies on the grating lobes due to non-plane-wave incidence are in progress and will provide a better understanding of the performance of the dichroic plate when illuminated by a corrugated horn.

### Acknowledgments

The author would like to thank P. Stanton and H. Reilly for performing the measurements, D. Hoppe and R. Hodges for technical discussions, and L. Epp for correcting a coding error in the program.

### References

- [1] T. Y. Otoshi and M. M. Franco, "Dual Passband Dichroic Plate for X-Band," *TDA Progress Report 42-94*, vol. April-June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 110-134, August 15, 1988.
- [2] N. Amitay, V. Galindo, and C. P. Wu, *Theory and Analysis of Phased Array Antennas*, New York: Wiley-Interscience, 1972.
- [3] C. C. Chen, "Transmission Through a Conducting Screen Perforated Periodically with Apertures," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, no. 9, pp. 627-623, September 1970.
- [4] C. C. Chen, "Transmission of Microwaves Through Perforated Flat Plates of Finite Thickness," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-21, no. 1, pp. 1-6, January 1973.
- [5] S. W. Lee, W. R. Jones, and J. J. Campbell, "Convergence of Numerical Solutions of Iris-Type Discontinuity Problems," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-19, no. 6, pp. 528-536, June 1971.
- [6] R. Mittra, T. Itoh, and T. S. Li, "Analytical and Numerical Studies of the Relative Convergence Phenomenon Arising in the Solution of an Integral Equation by the Moment Method," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, no. 2, pp. 96-104, February 1972.
- [7] S. W. Lee, G. Zarrillo, and C. L. Law, "Simple Formulas for Transmission Through Periodic Metal Grids or Plates," *IEEE Trans. Antenna Propagat.*, vol. AP-30, no. 5, pp. 904-909, September 1982.

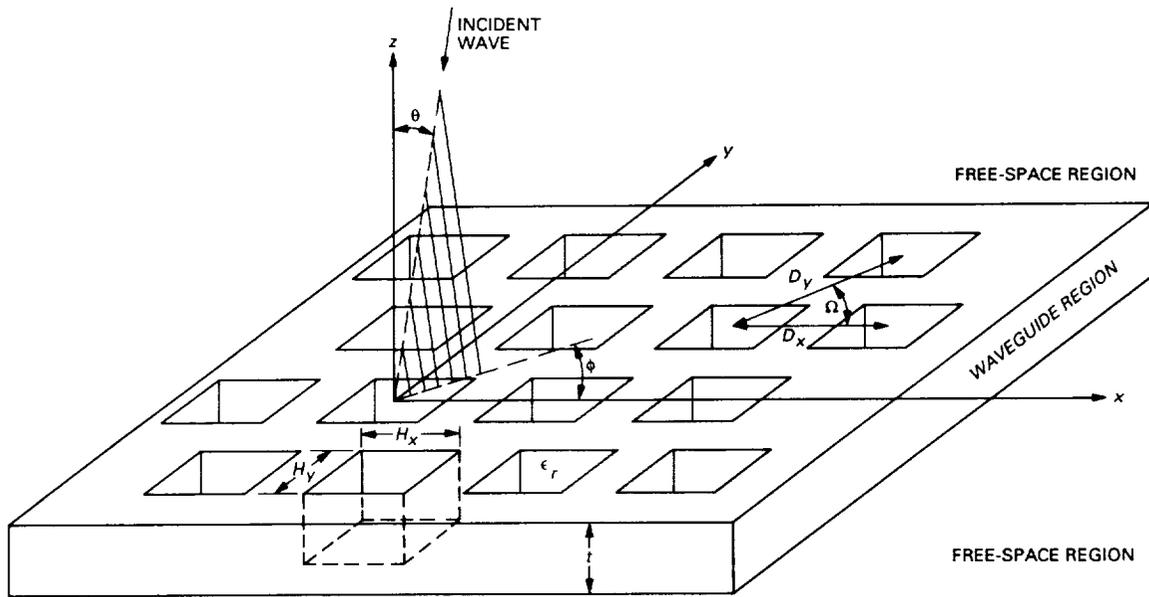


Fig. 1. Geometry of a thick dichroic plate with rectangular holes.

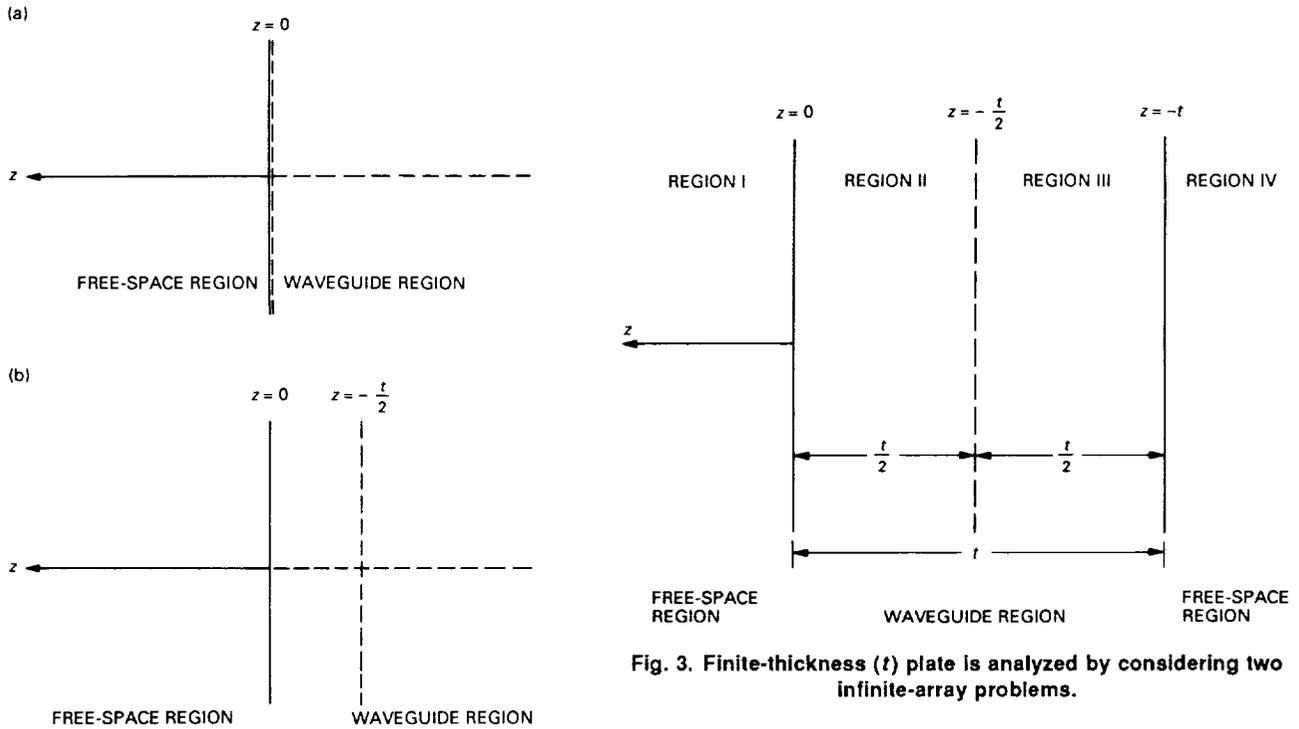


Fig. 2. Half-space infinite array with: (a) reference plane  $z = 0$ , and (b) reference plane  $z = -t/2$ .

Fig. 3. Finite-thickness ( $t$ ) plate is analyzed by considering two infinite-array problems.

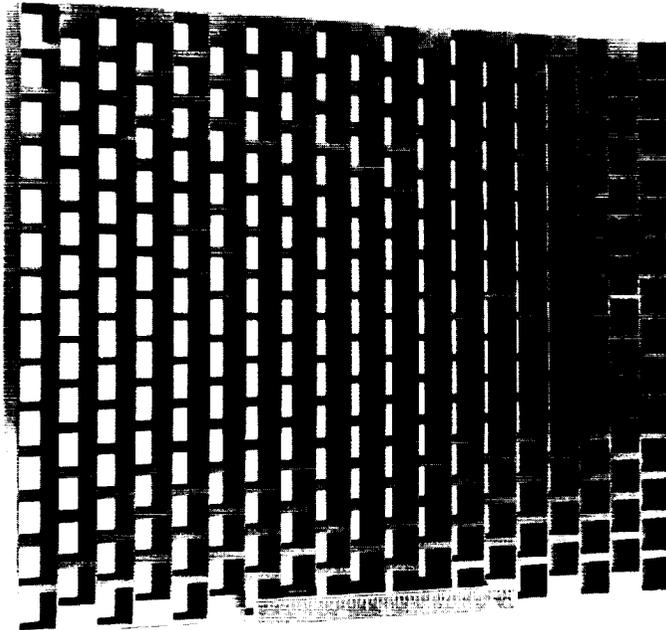


Fig. 4. Prototype dichroic plate.

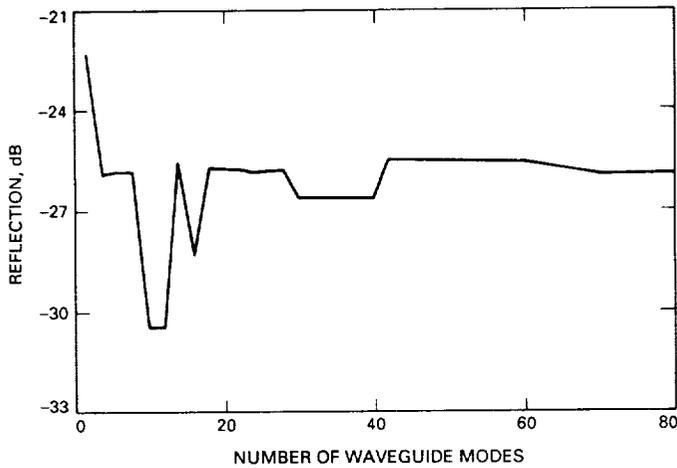


Fig. 5. Convergence with respect to the number of waveguide modes.

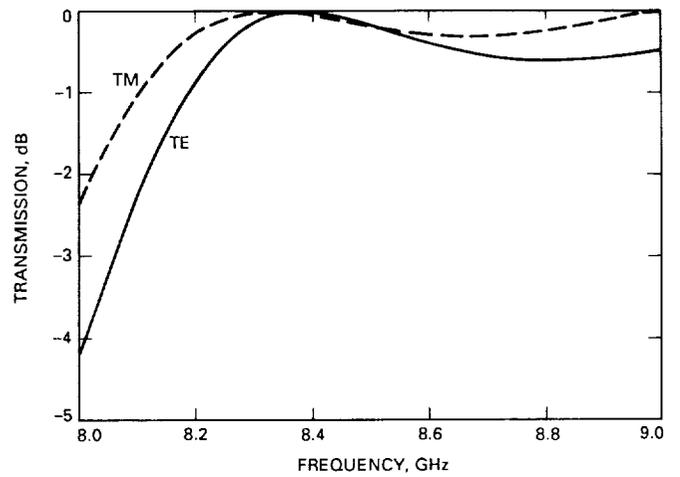


Fig. 6. Theoretical transmission versus frequency for the test dichroic plate for TE and TM polarization.

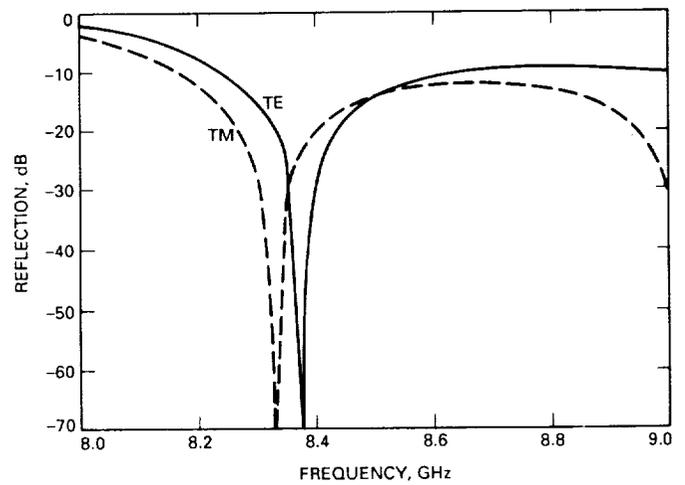


Fig. 7. Theoretical reflection versus frequency for the test dichroic plate for TE and TM polarization.

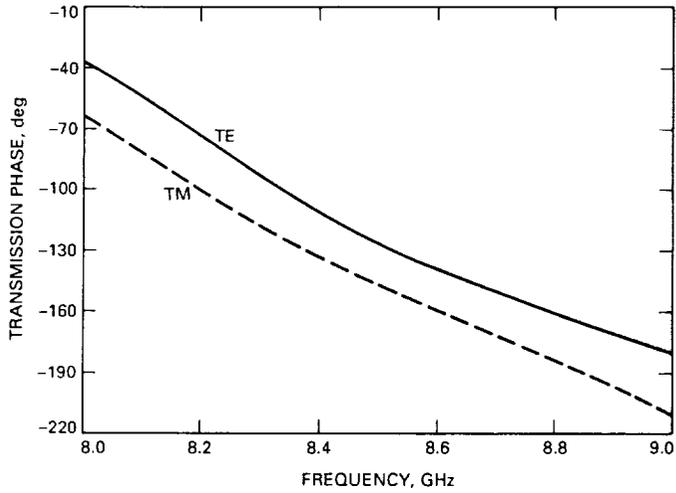


Fig. 8. Theoretical transmission phase versus frequency for the test dichroic plate for TE and TM polarization.

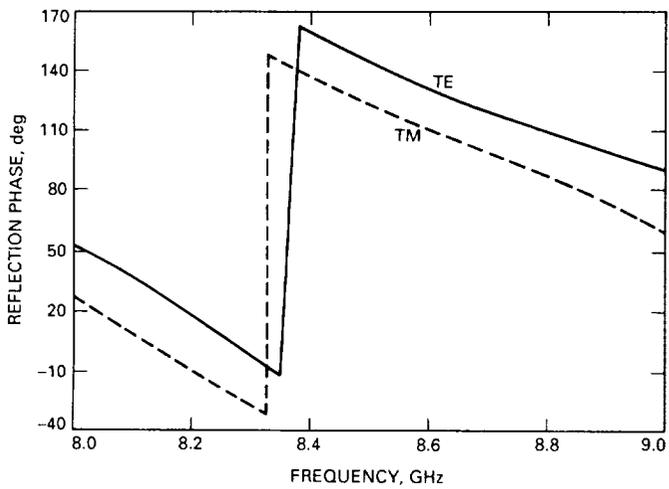


Fig. 9. Theoretical reflection phase versus frequency for the test dichroic plate for TE and TM polarization.

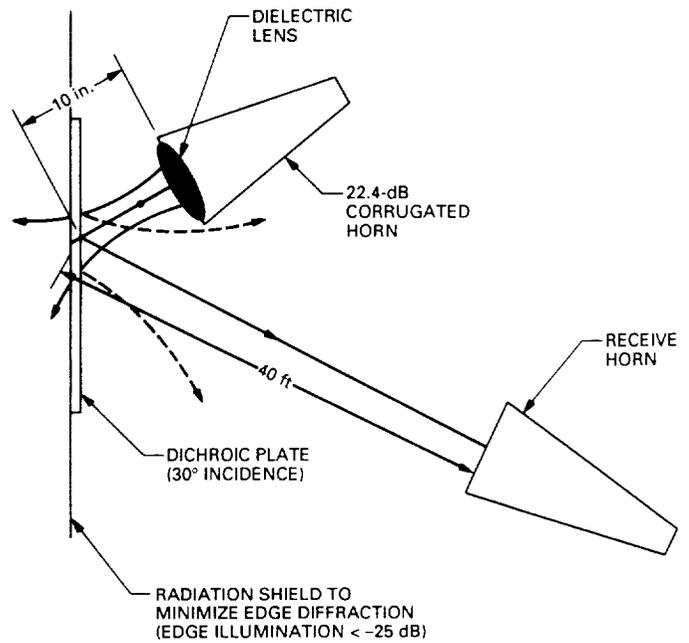


Fig. 10. Test setup for reflection measurements.

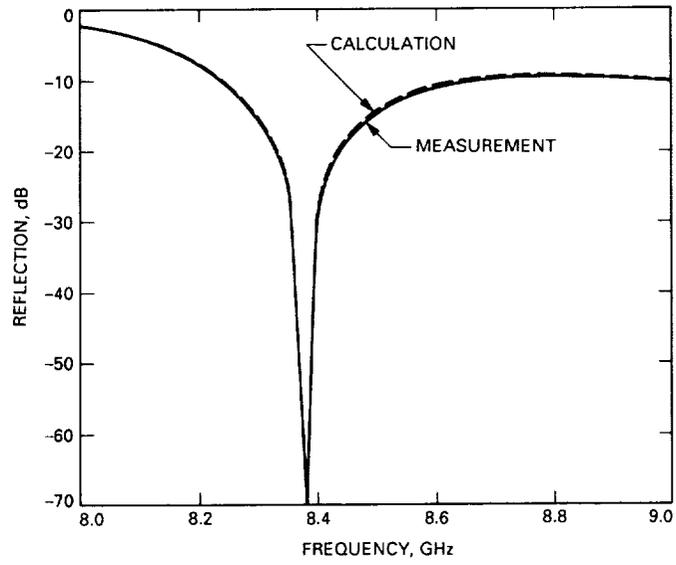


Fig. 11. Measured and calculated reflection versus frequency for the test dichroic plate for TE polarization.

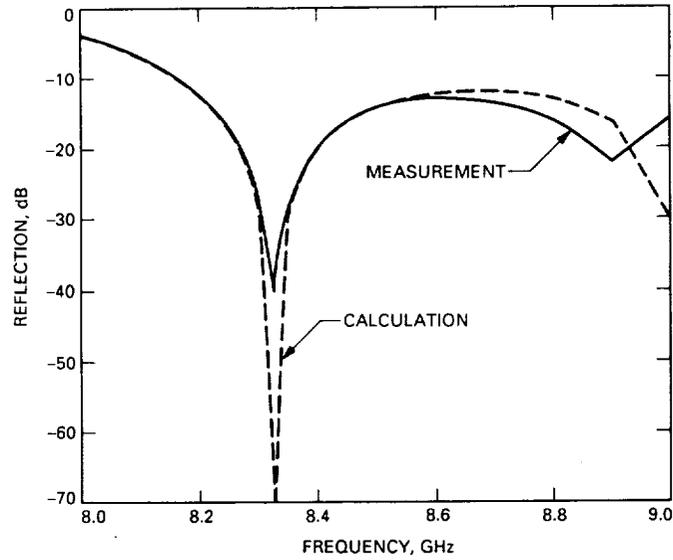


Fig. 12. Measured and calculated reflection versus frequency for the test dichroic plate for TM polarization.

53-32  
 P-13  
 N91-18311  
 JJ 57495

# An Analysis of the Least-Squares Problem for the DSN Systematic Pointing Error Model

L. S. Alvarez

Ground Antennas and Facilities Engineering Section

*A systematic pointing error model is used to calibrate antennas in the Deep Space Network. This article describes and analyzes the least-squares problem and the solution methods used to determine the model's parameters. Specifically studied are the rank-degeneracy problems resulting from beam-pointing error measurement sets that incorporate inadequate sky coverage. A least-squares parameter subset selection method is described and its applicability to the systematic error modeling process is demonstrated on a Voyager 2 measurement distribution.*

## I. Introduction

A pointing error model is used in the Deep Space Network's (DSN's) antenna-calibration process. With the exception of environmental effects, the major sources of errors in an antenna-pointing system are systematic and repetitive and therefore can be closely modeled. Examples of parameters in the model are residual errors in the geometric alignment of the mount axes and fixed-angle encoder offsets. Data collected from spacecraft and radio star observations are used to determine the parameters in the model and are then entered into the pointing system to accurately point the antenna. The origins of the pointing error modeling approach for radio-frequency (RF) antennas can be found in [1,2] while its development within the DSN is discussed in [3].<sup>1</sup>

The complete pointing error model is the sum of its separate error components. Table 1 shows individual error sources and the elevation/cross-elevation (or declination/cross-declination, depending on antenna mount) regressor variables used to estimate parameters. See [1,2,4] for a more thorough discussion of these parameters. Currently, this entire model is set in motion in the antenna-pointing system by entering parameter values manually. The DSN 70-m antennas track targets in both the computer command and precision modes of operation, each defined by a set of axis position transducers. (See [5] for a discussion of the axis servos and controllers.) The 34-m antennas employ only the computer command mode. In [3,6] recommended model parameter sets are given that apply to each tracking mode of these antennas; they are also repeated in Table 2. As can be seen, nine error parameters are used to estimate in the precision mode and eight in the computer command mode. In practice, the model parameters are determined by performing a least-squares fit

<sup>1</sup> R. L. Riggs, "Antenna Pointing Angle Corrections," DSN Antenna Seminar, Videotapes 49-54, Jet Propulsion Laboratory, Pasadena, California, May 1986.

on the pointing offset data collected from the spacecraft and/or from radio star observations. In this article, reference will only be made to the particular combinations of parameters in Table 2.

This article explores the numerical properties of the systematic error modeling process. Specifically, the analysis focuses on the numerical properties of the matrix formed by the pointing model regressor variables evaluated over the beam-pointing error data sets. These measurement sets may not cover enough points in the sky to accurately estimate all of the parameters. This is due to the finite number of targets and to other practical operational constraints, such as lack of antenna time. On the other hand, the objective of particular calibrations may be to optimize pointing in a particular region of the sky, such as along a constant declination. In practice, however, the limited measurement sets lead to rank deficiency in the least-squares measurement distribution matrix. This study of the problem will lead to a more objective approach to parameter selection and parameter estimate interpretation. In addition, the analytical techniques provided here may be used to predict which directions in the sky will yield optimal estimation.

The remainder of this article will formulate the systematic error parameter estimation problem and then establish a hypothetical performance index for matrix conditioning. In addition, the numerical tools presented will be used to analyze practical sky distributions in the context of the least-squares approximation and the current solution method will be reviewed. The article concludes with a proposed algorithm for parameter selection.

## II. Model Generation

In order to accurately point the antenna, pointing error correction models must be generated from radio star or spacecraft pointing offset data. This section deals with the model-fitting process, which uses the least-squares algorithm and assumes that the measurement data sets are accurate. At this time, the estimation process does not deal with uncertainties in the conical scan pointing offsets and radio star boresights except in human filtering of very large nonrepeatable and unexplainable offsets.

### A. Least-Squares Problem Formulation

The parameter vector  $\mathbf{p}$  of the systematic pointing error model is determined by performing a linear least-squares fit on the offset data. The estimation problem is formulated from  $m$  observations as

$$\begin{bmatrix} \delta x_{el_1} \\ \vdots \\ \delta x_{el_m} \\ \delta el_1 \\ \vdots \\ \delta el_m \end{bmatrix} = \begin{bmatrix} A_{xel_1}(el_1, az_1) \\ \vdots \\ A_{xel_m}(el_m, az_m) \\ A_{el_1}(el_1, az_1) \\ \vdots \\ A_{el_m}(el_m, az_m) \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} \quad (1)$$

or

$$\mathbf{y} = \mathbf{A}\mathbf{p} \quad (2)$$

where the offset vector  $\mathbf{y}$  is  $2m \times 1$ , the measurement distribution matrix  $\mathbf{A}$  is  $2m \times n$ , and the parameter vector  $\mathbf{p}$  is  $n \times 1$ . As can be seen, equations representing both the cross-elevation and elevation error functions are obtained for each single observation point in the sky. Let the least-squares estimator be  $\hat{\mathbf{p}}$  and satisfy the following matrix equation

$$\hat{\mathbf{y}} = \mathbf{A}\hat{\mathbf{p}} \quad (3)$$

where the vector  $\hat{\mathbf{y}}$  contains the estimated (or fitted) values to the cross-elevation and elevation offsets of Eq. (1). The difference between the individual elements, or residuals, is defined as

$$r_i = y_i - \hat{y}_i \quad (4)$$

The method of least squares chooses the parameter estimate  $\hat{\mathbf{p}}$ , such that the following quantity is minimized

$$\sum_{i=1}^{2m} r_i^2 \quad (5)$$

The estimate satisfies the following matrix equation [7,9]

$$\hat{\mathbf{p}} = (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{y} \quad (6)$$

where  $\mathbf{A}^t$  is the transpose of  $\mathbf{A}$ . Caution must be given to least-squares problems in which the regressor variables, or basis terms of  $\mathbf{A}$ , are not truly independent. In such cases the measurement distribution matrix  $\mathbf{A}$  may be close to, or is, rank deficient. If  $\mathbf{A}$  is rank deficient, then there are an infinite number of solutions to the least-squares problem and no conclusion can be drawn as to the role of the individual regressor variables [9].

During the systematic pointing error estimation process, limited data sets and inherent correlations in the pointing error model have led to rank deficiency and its associated problems. This situation was discussed in [8] where the condition of empirically correlated regressor variables was termed “multicollinearity.” It was pointed out in [8] that regressor variables of the model are not truly independent. However, this is not accurate when a proper combination of parameters is selected, as recommended in Table 2. The degree of linear independence in the columns of matrix  $\mathbf{A}$  for various antenna configurations is strongly dependent on the distribution of the observation points over the sky. This situation and its effect on the pointing error estimation is discussed below.

## B. All-Sky Model Analysis

An analytical approach was taken to obtain a performance index for the numerical conditioning of the systematic error least-squares problem and to compare it with results from practical measurement sets. One such performance index can be determined by examining a hypothetical all-sky uniform distribution of pointing offset data. These measurement points are used to generate measurement distribution matrices for different combinations of parameters. Intuitively, it would make sense to obtain pointing error offsets uniformly throughout the field of view of the antenna and conclude that this is the optimal distribution for input into the parameter estimation problem. However, observing the basis terms of the pointing model given in Table 2, it can be seen that not all terms are simultaneously functions of both azimuth and elevation. This condition will tend to result in redundant column elements of  $\mathbf{A}$ ; thus, optimal matrix conditioning will most likely not be obtained with the all-sky distribution. However, as will be shown, all-sky matrices do have acceptable conditioning and can be used for a suitable performance index. Singular value decomposition (SVD) was used to analyze the linear independence of the columns of  $\mathbf{A}$  and is defined in the following theorem.

**Theorem 1.** Let  $\mathbf{A}$  be a real  $m \times n$  matrix with  $m \geq n$ . Then there is an orthogonal matrix  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m]$  of order  $m$  and an orthogonal matrix  $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$  of order  $n$  such that

$$\mathbf{U}^t \mathbf{A} \mathbf{V} = \begin{bmatrix} \boldsymbol{\Sigma} \\ 0 \end{bmatrix} \quad (7)$$

where

$$\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n) \quad (8)$$

and

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0 \quad (9)$$

The theorem is taken from [10] and the more general SVD is proven in [9]. The numbers  $\sigma_1, \sigma_2, \dots, \sigma_n$ , which are unique, are called the singular values of  $\mathbf{A}$ . The columns  $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m]$  of  $\mathbf{U}$  are called the left singular vectors of  $\mathbf{A}$ , and the columns  $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$  of  $\mathbf{V}$  are called the right singular vectors of  $\mathbf{A}$ . SVD is extremely useful in analyzing numerical rank deficiency because the singular values indicate how near  $\mathbf{A}$  is to a matrix of lower rank. The matrix  $\mathbf{A}$  has rank  $r$  if and only if

$$\sigma_r > 0 = \sigma_{r+1} \quad (10)$$

Mathematically speaking, the smallest singular value of  $\mathbf{A}$  is the 2-norm distance of  $\mathbf{A}$  to the set of all rank-deficient matrices [9]. The ratio of the largest to smallest singular value is termed the condition number of  $\mathbf{A}$ . This number quantifies the sensitivity of the least-squares solution  $\hat{\mathbf{p}}$  of Eq. (2). Large condition numbers indicate that relatively small changes in  $\mathbf{A}$  or the offset vector  $\mathbf{y}$  can induce large changes in the computed least-squares solution  $\hat{\mathbf{p}}$ . This is undesirable since parameter vector estimates computed from such ill-conditioned measurement distribution matrices can lead to erroneous pointing offset corrections that will be applied in future antenna tracks. The pointing model corresponding to the DSS 14 70-m antenna (i.e., latitude = 35.426) was used to generate the full  $\mathbf{A}$  matrix of relevant error parameters. The uniform distribution consists of 9-deg increments in elevation and 20-deg increments in azimuth. This full measurement set is illustrated in Fig. 1. The singular values of the  $\mathbf{A}$  matrices corresponding to the precision mode and computer command mode of operation are presented in Table 3.

As can be seen, the numerical conditioning for the least-squares problem resulting from this hypothetical all-sky distribution is well behaved. Both the precision mode and computer command mode parameter sets yield  $\mathbf{A}$  matrices with reasonably nonzero singular values and small condition numbers, implying full-column rank. Another quantity commonly considered in least-squares analysis is the correlation matrix derived from  $(\mathbf{A}^t \mathbf{A})^{-1}$  of Eq. (6), which is numerically shown in Table 4. The matrix  $(\mathbf{A}^t \mathbf{A})^{-1}$  is an estimate of the covariance matrix for the solution vector of the least-squares problem. Values near one in the correlation matrix indicate high pairwise correlation between the estimated parameters.

It is evident that such high correlation is implied between the first three parameters (az collimation, fixed az encoder offset, and az/el axis skew) of the computer command mode set. Evaluation of the basis terms corresponding to these parameters in the limited elevation range of 0 to 90 deg results in the pairwise correlation and cannot be avoided regardless of the azimuth distribution. The implication of inherent correlation to parameter estimate stability was investigated through Monte Carlo simulations. The empirical estimation covariance matrix was computed and found to be in very close agreement with the theoretical covariance matrix computed from  $(\mathbf{A}^t \mathbf{A})^{-1}$ , thus illustrating that stable parameter estimates will result from an all-sky distribution. It was noted in the simulations that individual estimates of the first three elements always varied in the same direction of magnitude, but that differences never exceeded the bounds predicted in the theoretical standard-deviation vector given by

$$\sigma_{\mathbf{p}} = \sum_{i=1}^n \frac{\mathbf{v}_i}{\sigma_i} \quad (11)$$

where the  $\mathbf{v}_i$  and  $\sigma_i$  are defined in Theorem 1. The above equation is obtained by solving for  $\mathbf{A}$  in Eq. (7), substituting it into  $(\mathbf{A}^t \mathbf{A})^{-1}$ , and then taking the square root of the diagonal of the resulting matrix.

The numerical conditioning of the least-squares estimation of antenna precision and computer mode systematic error parameter sets was evaluated above for an all-sky distribution. The resulting measurement distribution matrix for each mode of operation was found to have full rank, thus ensuring unique least-squares solutions for  $\hat{\mathbf{p}}$ . Also, the large values in the correlation matrix were not seen to degrade the stability of repeated parameter estimates. The linear dependence of the parameters implied by the correlation matrix is due to their mathematical definitions and selecting them simultaneously will not degrade the estimate of the measurement vector  $\mathbf{y}$ . Such rich offset distributions can never be obtained in practice, thus it is inevitable that poorer matrix conditioning will lead to least-squares estimates of poorer quality. As shown by this analysis, the singular values and condition number of the distribution matrix  $\mathbf{A}$  are key parameters in evaluating ill-conditioned least-squares problems.

### C. Reduced and Sparse Data Sets

Current practices dictate that systematic error models be generated from antenna-pointing error-correction data taken from as much of the sky as possible or from an area defined by one or two declination angles. The first is used

to generate all-sky pointing models, while the second computes model parameters applicable only in limited directions of the sky. Both situations typically diverge from the hypothetical all-sky example since the basis terms of the pointing model are evaluated in fewer, and perhaps more redundant, directions. Their effect on the least-squares estimation process will be illustrated with examples.

Figure 2 shows the sky trajectory for the Voyager 2 spacecraft. Conical-scan offset data collected at a declination of  $-22.5$  deg clearly represent only a small portion of the total sky measurement space. Tables 5 and 6 show the singular values, condition numbers, and theoretical standard deviations in millidegrees (mdeg) for the least-squares estimate using the  $\mathbf{A}$  matrices generated for precision and computer command operation. As implied in the tables, matrix condition deteriorates in both parameter sets because of reduced measurement space. The theoretical standard deviations of the all-sky parameter sets are shown in mdeg in Table 7. Comparison with those of Table 6 illustrate the degradation of the least-squares parameter estimation. In [8], least-squares parameter fits were done on Voyager 1 conical-scan data obtained from the DSS 14 64-m antenna. The results were parameters that were too large in magnitude to be realistic or practical and that were unstable on a day-to-day basis. It has been shown through SVD analysis that such ill conditioning of the systematic error least-squares problem can, in general, be inferred a priori for any constant declination measurement set.

Figure 3 shows sky distribution for a radio source bore-sight offset file taken at the DSS 13 26-m antenna. The distribution is typical of data gathered during planetary radio astronomy experiments—here, for four radio sources. The pointing model regressor values were once again evaluated at the source coordinates and results of the SVD analysis are shown in Tables 8 and 9. Condition numbers for precision and computer command mode parameter sets are comparably small in magnitude to those from an all-sky distribution. The smallest singular values are also reasonably nonzero. Only minimal estimate degradation is predicted by the increase in theoretical standard deviations. Furthermore, the magnitude of this uncertainty is still reasonably small in the context of parameter estimates, which are usually in the tens of mdeg. This example illustrates that rank deficient measurement distribution matrices can be avoided by using recommended parameter sets and by evaluating the regressor variables with adequate sky distribution of pointing offsets. Concluding that such a measurement set is adequate for the least-squares model fitting is essentially putting emphasis on the norm of the resulting solution vector  $\hat{\mathbf{p}}$  instead of minimizing the norm of the

residual vector  $\mathbf{r}$  of Eq. (4). This approach appears to be the most logical given that the measurement uncertainties are not modeled. It has also been shown that, for radio source pointing calibrations, this matrix condition analysis can be done during pretrack activities, thus influencing the scheduling of calibrators.

It must be stressed that these results hold only for the sets of parameters recommended in Table 2. Different combinations of 21 error coefficients in the current pointing model will yield different, and in some cases disastrous, numerical properties of the matrices involved in the computation of the least-squares solution.

### III. Solution Methods

#### A. Parameter Selection

The two goals of the modeling process are to quantify contributors to the antenna's systematic pointing error so that pointing can be corrected and so that knowledge of the antenna's mechanical and structural characteristics can be acquired. To achieve both of these objectives simultaneously, identical parameter vectors must be chosen for estimation on a consistent basis. These parameters for the 70-m and 34-m antennas in the applicable mode of operation have been given in Table 2. Subsets of these vectors should be chosen either when parameter values are physically known a priori or when they are consistently estimated with small magnitudes. In practice, however, the goal of correct pointing can be achieved without accurate knowledge of actual antenna error characteristics. Optimization may be based on any random set of parameters that minimizes the sum of the squares of the residuals given in Eq. (5) without regard for physical interpretation.

Regardless of the estimation philosophy practiced, problems always arise when building models for particular regions of the sky—for example, along a line or band of constant declination for one or more sources—or for a particular spacecraft. The rank deficiency that plagues least-squares problems in these cases generates uncertainty in parameter selection and interpretation. However, such models for locally optimized pointing are needed for critical spacecraft and holography tracks and for those tracks of single sources known as strong, reliable antenna calibrators. The current least-squares solution method described next uses the SVD to accommodate ill-conditioned measurement distribution matrices.

#### B. Singular Value Decomposition

The SVD subroutines in the systematic error modeling software that were used to solve the least-squares problem

were taken from [11]. A key feature of the SVD method is its ability to handle rank deficiency. Ill-conditioned  $\mathbf{A}$  matrices result in the rank  $(\mathbf{A}) = r$  being less than the parameter dimension  $n$ . This results in a rank-deficient least-squares problem that has an infinite number of solutions, for if the vector  $\mathbf{p}$  is a minimizer and the vector  $\mathbf{z} \in \text{null}(\mathbf{A})$ , then  $\mathbf{p} + \mathbf{z}$  is also a minimizer. The SVD method is useful in such situations since it is a revealing and complete orthogonal decomposition. The routines from [11] basically implement the following theorem taken from [9], given here without proof.

**Theorem 2.** Suppose  $\mathbf{U}^t \mathbf{A} \mathbf{V} = \mathbf{\Sigma}$  is the SVD of  $\mathbf{A} \in \mathbb{R}^{m \times n}$  with  $r = \text{rank}(\mathbf{A})$ . If  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m]$  and  $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$  are column partitionings and  $\mathbf{y} \in \mathbb{R}^m$  then

$$\mathbf{p}_{\text{LS}} = \sum_{i=1}^r \frac{\mathbf{u}_i^t \mathbf{y}}{\sigma_i} \mathbf{v}_i \quad (12)$$

minimizes  $\|\mathbf{A}\mathbf{p} - \mathbf{y}\|_2$  and has the smallest 2-norm of all minimizers. Moreover

$$\|\mathbf{A}\mathbf{p}_{\text{LS}} - \mathbf{y}\|_2^2 = \sum_{i=r+1}^m (\mathbf{u}_i^t \mathbf{y})^2 \quad (13)$$

Note that if  $r < n$ , this corresponds to simply adding a zero multiple to the solution vector  $\mathbf{p}_{\text{LS}}$  rather than adding random large-valued multiples produced by the near-zero singular values. This may reduce uncertainty in the estimated coefficients, as in Eq. (11), but increases the residual norm, as in the increased summation index of Eq. (13). This point was touched on earlier. In practice, one must still come up with a numerical estimate  $\hat{r}$  of  $r$ . The systematic error modeling software estimates the numerical rank  $\hat{r}$  of  $\mathbf{A}$  as

$$\sigma_1 \geq \dots \geq \sigma_{\hat{r}} \geq \delta > \sigma_{\hat{r}+1} \geq \dots \geq \sigma_n \quad (14)$$

where the tolerance  $\delta$  is chosen to be  $\sigma_1$ , scaled by a machine-precision dependent factor. The selection of  $\hat{r}$  results in

$$\mathbf{p}_{\hat{r}} = \sum_{i=1}^{\hat{r}} \frac{\mathbf{u}_i^t \mathbf{y}}{\sigma_i} \mathbf{v}_i \quad (15)$$

as an approximation to  $\mathbf{p}_{\text{LS}}$ . If  $\sigma_{\hat{r}} \gg \delta$ , then  $\mathbf{p}_{\hat{r}}$  is a very close approximation to the true minimizer  $\mathbf{p}_{\text{LS}}$  since

$\mathbf{A}$  can be unambiguously regarded as a matrix with rank  $\hat{r}$  [9]. When  $[\sigma_1, \dots, \sigma_n]$  do not clearly split into small and large values, rank determination may be somewhat arbitrary.

### C. A New Algorithm for Parameter Selection

As has been shown, the SVD solution currently alleviates the rank-deficiency problems associated with limited pointing offset distributions. This means filtering out small singular values of  $\mathbf{A}$  and replacing them with the matrix  $\mathbf{A}_{\hat{r}}$  defined as

$$\mathbf{A}_{\hat{r}} = \sum_{i=1}^{\hat{r}} \sigma_i \mathbf{u}_i \mathbf{v}_i^t \quad (16)$$

where  $\hat{r}$  is the numerically determined estimate of the rank of  $\mathbf{A}$ . As discussed in [9] such a cutoff makes sense when the measurement distribution matrix is derived from noisy data. However, in this case,  $\mathbf{A}$  is being evaluated using accurate ephemeris from observed targets, as in Eq. (1). In other applications, rank deficiency is an indication of redundancy among factors that comprise the model. As has been shown in previous sections, redundancy among systematic error regressor variables occurs in the estimation process only when dealing with limited and reduced pointing measurement sets. In these cases, the systematic error predictor  $\mathbf{A}_{\hat{r}} \mathbf{p}_{\hat{r}}$  used in subsequent tracking will involve all  $n$  redundant factors that may have been chosen as a result of random parameter selection. Although such solutions may correct future pointing, parameter estimates can obscure physical interpretation of true antenna mechanical characteristics. In such instances, it is argued that the least-squares solution vector should contain at most  $\hat{r}$  nonzero systematic error parameters, which in turn dictate which columns of  $\mathbf{A}$  will be used in approximating the observation vector  $\mathbf{y}$ . The problem of choosing the appropriate columns of the measurement distribution matrix is termed subset selection. The SVD-based subset selection procedure that has been chosen for this least-squares application is summarized below:

- (1) Compute the SVD  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^t$  and use it to determine a rank estimate  $\hat{r}$ .
- (2) Calculate a permutation matrix  $\mathbf{P}$  such that the columns of the matrix  $\mathbf{B}_1 \in \mathbb{R}^{m \times \hat{r}}$  in  $\mathbf{A}\mathbf{P} = [\mathbf{B}_1 \ \mathbf{B}_2]$  are “sufficiently independent.”
- (3) Predict  $\mathbf{y}$  with the vector  $\mathbf{A}\mathbf{p}_{\text{sub}}$  where  $\mathbf{p}_{\text{sub}} = [\mathbf{z} \ \mathbf{0}]^t$  and  $\mathbf{z} \in \mathbb{R}^{\hat{r}}$  minimizes  $\|\mathbf{B}_1 \mathbf{z} - \mathbf{y}\|_2$ .

Using systematic error modeling, the rank determination in the first step can be chosen with more heuristic

criteria instead of those used in Eq. (14). The new criteria are based on the matrix condition number and the magnitude of the theoretical standard-deviation vector given by Eq. (11). Given  $\hat{r}$ , the first  $\hat{r}$  columns of permutation matrix  $\mathbf{P}$  give the column indices of  $\mathbf{A}$  for use in the least-squares estimation. These are equivalent to the parameters from which the model is selected. A thorough discussion of the various approaches to this problem can be found in [9,10]. Below is a summary of the algorithm to compute  $\mathbf{P}$  that was chosen and implemented in the systematic error modeling software. It is based on both the SVD and on QR factorization with the column-pivoting algorithm. For  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , QR factorization with column pivoting from [9] produces  $\mathbf{A}\mathbf{P} = \mathbf{Q}\mathbf{R}$  where

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{matrix} \hat{r} \\ m - \hat{r} \\ \hat{r} & n - \hat{r} \end{matrix} \quad (17)$$

where  $\hat{r}$  is the rank ( $\mathbf{A}$ ),  $\mathbf{Q}$  is orthogonal,  $\mathbf{R}_{11}$  is upper triangular and nonsingular, and  $\mathbf{P}$  is a permutation matrix. This factorization implies that the first  $\hat{r}$  columns of  $\mathbf{Q}$  form an orthonormal basis for range ( $\mathbf{A}$ ). It is the desired result since the measurement vector  $\mathbf{y}$  in the least-squares problem may be approximated by the first  $\hat{r}$  columns of the matrix  $\mathbf{A}\mathbf{P}$ , which is just  $\mathbf{B}_1$  of the second step above. This is equivalent to choosing the first  $\hat{r}$  parameters of  $\mathbf{P}\mathbf{p}$  for estimation, which is equal to the vector  $\mathbf{z}$  in step three above. As in a previous section where the case for cutting off singular values in the SVD method was presented, reducing the order of the parameter solution vector will also increase the residual norm.

Unfortunately, QR factorization with column pivoting alone is not a totally robust method for computing the permutation matrix  $\mathbf{P}$  [9]. The preferred algorithm implemented in the software that uses both SVD and QR factorization is presented in the revised steps below:

- (1) Compute the SVD  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^t$  and use it to determine a rank estimate  $\hat{r}$ . Save the matrix  $\mathbf{V}$ .
- (2) Apply the QR factorization with column pivoting to the subset of  $\mathbf{V}^t$ :  $\mathbf{Q}^t \mathbf{V}^t(:, 1 : \hat{r})^t \mathbf{P} = [\mathbf{R}_{11} \ \mathbf{R}_{12}]$  and set  $\mathbf{A}\mathbf{P} = [\mathbf{B}_1 \ \mathbf{B}_2]$  with  $\mathbf{B}_1 \in \mathbb{R}^{m \times \hat{r}}$  and  $\mathbf{B}_2 \in \mathbb{R}^{m \times (n - \hat{r})}$ .
- (3) Determine  $\mathbf{z} \in \mathbb{R}^{\hat{r}}$ , which minimizes  $\|\mathbf{B}_1 \mathbf{z} - \mathbf{y}\|_2$ .

The main contribution of this algorithm is facilitating parameter selection for reduced and constant declination-pointing measurement sets. In the latter case, its application will ensure consistent parameter selection for particular radio sources and spacecraft tracks. This subset

selection procedure essentially eliminates parameters that the algorithm has deemed unobservable in the given measurement distribution. The next step is to decide how to deal with these excluded parameters. One approach is to simply ignore them and proceed as usual with the least-squares estimation with the reduced vector  $\mathbf{z}$ , as determined above.

A different approach, when possible, is to use physically known or accurate a priori estimates for unobservable parameters and subtract their contributions from the measurement vector  $\mathbf{y}$  before estimating  $\mathbf{z}$ . Such an option is available in the current software. The resultant solution vector should be more consistent with all-sky models. Finally, it should be noted that the tools presented here can be used for the opposite effect (e.g., predicting matrix condition and rank or for least-squares estimate accuracy) when the measurement vector  $\mathbf{y}$  is augmented with pointing offsets taken in new directions.

To illustrate, this algorithm is applied to the  $\mathbf{A}$  matrix which resulted from the Voyager 2 trajectory, as shown in Fig. 2. Referring to Table 5, one can base the rank determination of the  $\mathbf{A}$  matrix on the smallest of the singular values. For example, choosing 0.1 as a singular value cutoff results in precision and computer command mode parameter selections and matrix conditions that are summarized in Tables 10 and 11. Eliminating parameters 11 and 21 from the precision mode and 1 from the computer command mode results in reduced matrix condition and smaller estimation standard deviations for some elements of the solution vector. In practice, the actual systematic error estimated values are generally less than 100 mdeg. Thus, estimation accuracy for some of the remaining parameters in Table 10 will be a certain percentage of the estimated values.

Depending on the antenna's frequency band, this may or may not meet the pointing requirements. (A detailed description of errors will not be covered here.) Estimation errors will always be larger in practice because of uncertainties in the measurement vector  $\mathbf{y}$ , so one may decide to increase the singular value cutoff and apply the subset selection algorithm. Using cutoffs 1.0 for the precision mode and 0.2 for the computer command mode yields the results summarized in Tables 12 and 13. To achieve accuracy comparable to the hypothetical all-sky models, the parameters to be excluded are 1, 7, 11, and 21 from the precision-mode set and 1 and 7 from the computer command-mode set. It is advised that whenever the fixed angular encoder error parameters (for example, 7 and 21) are excluded in the subset selection procedure, their values should be determined directly from the pointing offset data

and contributions to  $\mathbf{y}$  should be removed before making an estimation.

## IV. Summary

This article has described and analyzed the least-squares problem encountered in the DSN systematic pointing error modeling process. Specifically investigated is the relationship between rank degeneracy of measurement distribution matrices and limited-sky distributions of the pointing error offsets. Using a hypothetical all-sky performance index and an SVD analysis, it is shown that an acceptable matrix condition of the least-squares problem can be obtained by evaluating the pointing model regressor variables with adequate sky distributions of the pointing measurements. In addition to matrix condition, the theoretical standard deviations of the least-squares estimate are used to evaluate accuracy. It is shown through an example that redundancy among the systematic error model regressor variables occurs when dealing with limited and sparse data sets. In practice, rank-degenerate matrices are encountered when building models for particular regions of the sky, such as along a band of constant declination.

The key feature of the analysis presented is its predictive capability. Matrix condition and least-squares estimate accuracy based on measurement distribution may be predicted before actual pointing calibration activities commence. The current least-squares solution method based on singular-value decomposition is also presented. This method can handle ill-conditioned measurement distribution matrices encountered in the model-building process. For limited measurement sets, it was argued that it may be preferred to estimate only observable parameters. Systematically eliminating redundant parameters will facilitate the parameter selection process and make it consistent. A recommended subset selection algorithm based on singular-value decomposition and QR factorization is illustrated with a Voyager 2 measurement set.

## V. Future Work

This article has presented an analytical approach using mathematical tools to answer fundamental numerical questions arising from the systematic error-modeling process. Such general but consistent procedures are needed in the modeling process because of the many antenna-specific mechanical and other practical considerations encompassed by the problem. Once past this juncture, one may begin to address the deficiencies and look for possible refinements in the estimation process. The most obvious is that of recursive estimation. Methods must be devised to handle

data sets spanning many weeks or years and incorporating many a priori models and model uncertainties into the estimation algorithms. If and when uncertainties in pointing measurements can be accurately modeled, including those from natural or manmade sources as well as from antenna-system imperfections, then the algorithm should also be modified to allow for weighted observations.

All these enhancements must be worked into the existing modeling software. This package should also enable the functional form of the model to change relatively often. This will allow for the addition of newly discovered error terms and for enhancements to accommodate new antenna architectures such as the DSS 13 beam-waveguide antenna.

## Acknowledgments

The author would like to thank R. Riggs for instruction about the DSN systematic pointing error modeling process. The modeling software developed and maintained by Riggs for the Goldstone tracking station was used in this study. Chris Sanelli is acknowledged for implementing and testing the parameter selection algorithm.

## References

- [1] P. Stumpff, "Astronomical Pointing Theory for Radio Telescopes," *Klein-Heibacher Berichte*, vol. 15, Formoldeteshnischon Zentralamt, Darmstadt, West Germany, pp. 431-437, 1972.
- [2] M. L. Meeks, H. A. Ball, and A. B. Hull, "The Pointing Calibration of the Haystack Antenna," *IEEE Transactions on Antennas and Propagation*, vol. AP-16, no. 6, pp. 746-751, November 1968.
- [3] R. Stevens, R. L. Riggs, and B. Wood, "Pointing Calibration of the MKIVA DSN Antennas for Voyager 2 Uranus Encounter Operations Support," *TDA Progress Report 42-87*, vol. July-September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 206-239, November 15, 1986.
- [4] C. N. Guiar, F. L. Lansing, and R. Riggs, "Antenna Pointing Systematic Error Model Derivations," *TDA Progress Report 42-88*, vol. October-December 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 36-46, February 15, 1987.
- [5] R. E. Hill, "A Modern Control Theory Based Algorithm for Control of the NASA/JPL 70-Meter Antenna Axis Servos," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 285-294, November 15, 1987.
- [6] C. N. Guiar, R. L. Riggs, R. Stevens, and M. Wert, "DSS 14 Antenna Calibrations for GSSR/VLA Saturn Radar Experiments," *TDA Progress Report 42-93*, vol. January-March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 309-337, May 15, 1988.
- [7] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1974.

- [8] C. N. Guiar, "Using Ridge Regression in Systematic Error Corrections," *TDA Progress Report 42-92*, vol. October–December 1987, Jet Propulsion Laboratory, Pasadena, California, February 15, 1988.
- [9] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Baltimore, Maryland: The Johns Hopkins University Press, 1989.
- [10] G. H. Golub, V. Klema, and G. W. Stewart, *Rank Degeneracy and Least-Squares Problems*, Technical Report TR-456, Department of Computer Science, University of Maryland, College Park, Maryland, 1976.
- [11] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge, England: Cambridge University Press, 1986.

**Table 1. Systematic pointing error sources and model terms**

Error source	Model function	
	Cross-elevation error	Elevation error
Az <sup>a</sup> collimation	$P_1^e$	-
Az encoder fixed offset	$P_2 \cos(\text{el})$	-
Az/el skew	$P_3 \sin(\text{el})$	-
Az axis tilt	$P_4 \sin(\text{el}) \cos(\text{az})$	$-P_4 \sin(\text{az})$
Az axis tilt	$P_5 \sin(\text{el}) \sin(\text{az})$	$P_5 \cos(\text{az})$
Source dec <sup>b</sup>	$P_6 \sin(\text{az})$	$P_6 \sin(\text{el}) \cos(\text{az})$
El <sup>c</sup> encoder fixed offset	-	$P_7$
Gravitational flexure	-	$P_8 \cos(\text{el})$
Residual refraction	-	$P_9 \cot(\text{el})$
Az encoder scale error	$P_{10} (\text{az}/360) \cos(\text{el})$	-
	Cross-declination error	Declination error
HA <sup>d</sup> /dec axis skew	$-P_{11} \sin(\text{dec})$	-
HA axis tilt	$P_{12} \sin(\text{HA}) \sin(\text{dec})$	$P_{12} \cos(\text{HA})$
HA axis tilt	$-P_{13} \cos(\text{HA}) \sin(\text{dec})$	$P_{13} \sin(\text{HA})$
HA feed offset	$-P_{14}$	-
Gravitational flexure	$P_{15} \cos(p)^f \cos(\text{el})$	$-P_{15} \sin(p) \cos(\text{el})$
Declination feed offset	-	$P_{16}$
Gravitational flexure	$P_{17} \sin(p) \cos(\text{el})$	-
Gravitational flexure	-	$-P_{18} \cos(p) \cos(\text{el})$
Gravitational flexure	$-P_{19} \sin(\text{el})$	-
Gravitational flexure	-	$P_{20} \sin(\text{el}) (\text{el})$
HA encoder bias	$P_{21} \cos(\text{dec})$	-

<sup>a</sup> Az refers to azimuth angle.      <sup>d</sup> HA refers to hour angle.  
<sup>b</sup> Dec refers to declination angle.    <sup>e</sup> Uppercase *P* refers to parameter value.  
<sup>c</sup> El refers to elevation angle.        <sup>f</sup> Lowercase *p* refers to paralectic angle.

**Table 2. Applicable parameter sets to DSN 70-m and 34-m antennas**

Precision mode	Computer command mode
1	1
7	2
8	3
9	4
11	5
12	7
13	8
14	9
21	-

**Table 3. Singular values for all-sky distributions**

Precision mode	Computer command mode
34.252	34.252
18.108	18.096
15.785	11.814
11.811	11.807
10.342	10.342
6.161	5.644
4.131	2.623
2.925	0.819
2.623	-
Condition number	Condition number
13.06	41.80

**Table 4. Computer command mode correlation matrix for all-sky distribution**

Parameter	1	2	3	4	5	7	8	9
1	1.00	-0.97	-0.98	0.00	0.00	0.00	0.00	0.00
2	-0.97	1.00	0.92	0.00	0.00	0.00	0.00	0.00
3	-0.98	0.92	1.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	1.00	-0.81	0.36
8	0.00	0.00	0.00	0.00	0.00	-0.81	1.00	-0.75
9	0.00	0.00	0.00	0.00	0.00	0.36	-0.75	1.00

**Table 5. Singular values for the Voyager 2 data set**

Precision mode	Computer command mode
57.226	57.226
14.530	12.171
10.282	10.282
5.912	5.042
1.483	1.483
0.334	1.315
0.199	0.199
0.0033	0.0671
0.0028	-
Condition number	Condition number
20,438	853

**Table 7. Theoretical standard deviations for the all-sky data set in millidegrees**

Parameter	Precision mode	Computer command mode
1	0.16	0.82
2	-	0.61
3	-	0.69
4	-	0.08
5	-	0.08
7	0.17	0.17
8	0.35	0.35
9	0.06	0.06
11	0.23	-
12	0.08	-
13	0.21	-
14	0.20	-
21	0.22	-

**Table 6. Theoretical standard deviations for the Voyager 2 data set in millidegrees**

Parameter	Precision mode	Computer command mode
1	2.19	10.46
2	-	10.06
3	-	3.46
4	-	0.20
5	-	0.72
7	3.55	3.55
8	3.53	3.53
9	0.03	0.03
11	289.48	-
12	0.72	-
13	1.89	-
14	252.03	-
21	269.20	-

**Table 8. Singular values for radio source distribution**

Precision mode	Computer command mode
15.951	15.452
13.006	13.157
9.885	8.176
7.660	7.504
4.326	4.238
3.550	2.603
2.020	0.826
1.156	0.283
0.825	-
Condition number	Condition number
19.335	54.600

**Table 9. Theoretical standard deviations for the radio source data set in millidegrees**

Parameter	Precision mode	Computer command mode
1	0.33	2.45
2	—	1.68
3	—	1.96
4	—	0.13
5	—	0.14
7	0.42	0.42
8	1.11	1.11
9	0.36	0.35
11	0.44	—
12	0.14	—
13	0.44	—
14	0.49	—
21	0.57	—

**Table 10. Reduced parameter standard deviations for the Voyager 2 data set in millidegrees**

Parameter	Precision mode	Computer command mode
1	2.16	—
2	—	0.29
3	—	0.72
4	—	0.20
5	—	0.72
7	3.55	3.55
8	3.52	3.53
9	0.03	0.03
12	0.72	—
13	1.87	—
14	1.30	—

**Table 11. Singular values for reduced Voyager 2 parameter set**

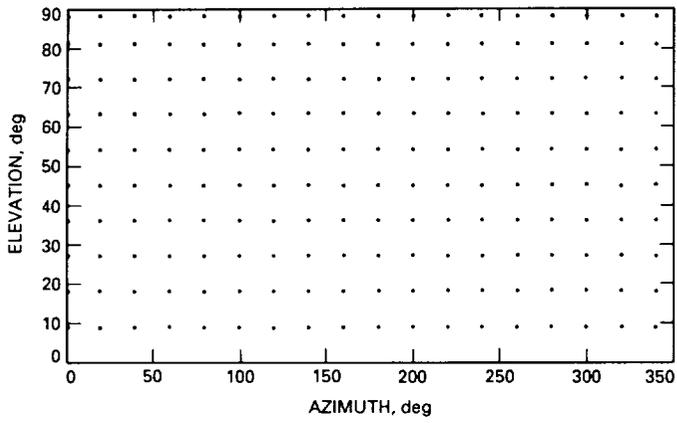
Precision mode	Computer command mode
57.226	57.226
12.157	10.282
10.282	8.893
5.221	4.845
1.483	1.483
0.319	1.315
0.199	0.199
Condition number	Condition number
288	288

**Table 12. Reduced parameter standard deviations for the Voyager 2 data set in millidegrees**

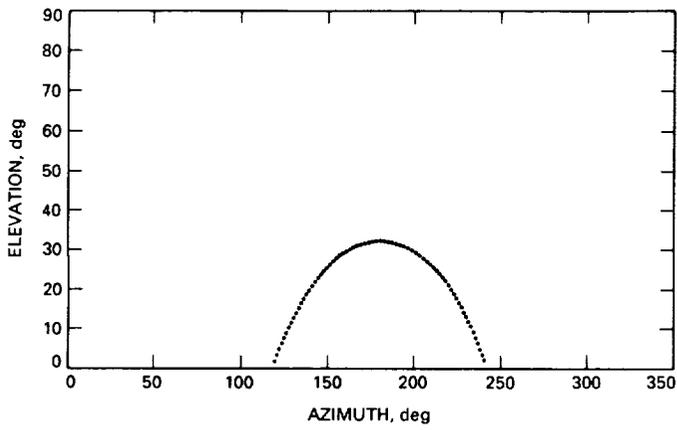
Parameter	Precision mode	Computer command mode
2	—	0.29
3	—	0.72
4	—	0.20
5	—	0.53
7	—	—
8	0.55	0.55
9	0.03	0.03
12	0.53	—
13	0.19	—
14	0.13	—

**Table 13. Singular values for reduced Voyager 2 parameter set**

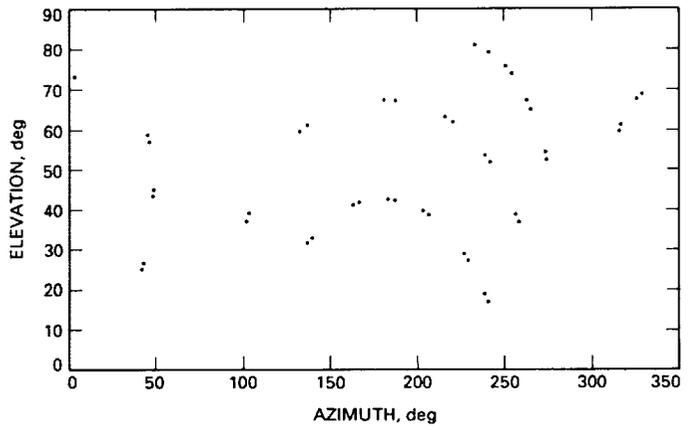
Precision mode	Computer command mode
56.955	56.955
8.952	8.893
8.137	8.137
4.900	4.845
1.335	1.335
—	1.315
Condition number	Condition number
43	43



**Fig. 1. Hypothetical all-sky measurement set.**



**Fig. 2. Voyager 2 measurement set.**



**Fig. 3. Radio source boresight measurement set.**

54-32

330062  
V/20

**N91-18312**

JJ 70450

# Predictive Control and Estimation Algorithms for the NASA/JPL 70-Meter Antennas

W. Gawronski

Ground Antennas and Facilities Engineering Section

*This article presents a modified output-prediction procedure and a new controller design based on the predictive control law. Also, a new predictive estimator is developed to complement the controller and to enhance system performance. The predictive controller is designed and applied to the tracking control of the Deep Space Network 70-m antennas. Simulation results show significant improvement in tracking performance over the linear-quadratic controller and estimator presently in use.*

## I. Introduction

The recent pointing requirements for the X-band (8.4 GHz) frequency, 70-m aperture antenna (further denoted as the DSS 14 antenna) and the expectation of future K-band (32 GHz) capability dictate a need for high-performance controllers for the azimuth and elevation drives. This article presents a new design procedure for a tracking controller that significantly improves antenna tracking performance. It considers on-axis (or servo) tracking. In this case, the output is taken on the encoder (or tachometer) rather than the radio frequency (rf) pointing position. The predictive controller uses future values of the stored input command to generate the control signal. For this reason, the predictive control principle is considered useful in design of the Deep Space Network (DSN) antennas' tracking controllers, since the antenna future tracking command is known when following stars or spacecraft.

The tracking-control problem is a nontrivial extension of the regulator problem, widely investigated in the

control literature. Several approaches to the solution of the tracking problem have been presented in [1-11]. Predictive controllers are described and analyzed in many papers, among them [1,4,5,6,7,9,10,11]. In most of them, controlled auto-regressive and integrated moving-average (CARIMA) models are developed and extensively used for output prediction and predictive control of linear systems [1,4,5,6,7,9,10]. State-space description serves as a standard tool for system analysis and design. Besides tool standardization, the state-space representation of a predictive control system provides a unique insight into system properties, improves system design, and simplifies analysis. Interpretations of CARIMA modeling in the state space are provided by [1,4,11].

This article presents a state-space description of the output-prediction procedure and a new controller design based on the predictive control law. A new input-reference scheme that uses the input-reference horizon is introduced. Thus, the increment of the control signal is determined with respect to the input horizon rather than to the last

value of the input. Also, the introduced weighting matrix includes a forgetting factor. Both features significantly improve system performance.

The usefulness of the predictive control law depends on the availability of the plant-state variables for measurement. Typically, not all state variables can be measured, although for an observable system they can be estimated. In this article, a predictive estimator is developed as a complement to the predictive controller to speed the estimation process and enhance system performance.

Using the introduced predictive control and estimation laws, the state-space predictive controller is designed and applied to the tracking control of the DSN 70-m antennas. Simulation results show significant improvement in tracking performance over the linear-quadratic (LQ) controller and estimator presently in use. The robustness to parameter variation and the disturbance-suppression properties are found to be fairly good for the considered predictive control system.

## II. Output Prediction for a Linear System

A plant with  $nu$  inputs and  $ny$  outputs is considered. Its linear state-space model consists of  $n$  states

$$\mathbf{x}(i+1) = A\mathbf{x}(i) + B\mathbf{u}(i), \quad \mathbf{y}(i) = C\mathbf{x}(i) + D\mathbf{u}(i) \quad (1)$$

where  $\mathbf{x} \in R^n$ ,  $\mathbf{u} \in R^{nu}$ , and  $\mathbf{y} \in R^{ny}$ . The task is to predict output  $\mathbf{y}$  for  $NY$  steps ahead, given projected input  $\mathbf{u}$  for  $NU$  steps ahead. The integer  $NY$  is the length of the output horizon, while  $NU$  is the length of the input horizon. For casual systems, the length of the input horizon does not exceed the length of the output horizon, i.e.,  $NU \leq NY$ .

Before determining the predicted output, the input and output sequences (or horizons) are introduced. Three types of input sequences are defined. First is the input horizon  $U(i)$ , consisting of the input from an instant  $i$  up to  $NU - 1$  steps ahead

$$U^T(i) = [u_i^T(0), u_i^T(1), \dots, u_i^T(NU - 1)] \quad (2)$$

where  $u_i(k)$  is the predicted input at instant  $i$  with  $k$  steps ahead. The input horizon  $U(i-1)$  is a horizon predicted at the previous time instant. Note that it is not a delayed prediction at instant  $i$ , i.e.,  $u_{i-1}(k) \neq u_i(k-1)$ .

The second sequence, reference-input horizon  $U_r(i)$ , is identical to the previous input horizon  $U(i-1)$  for the

first  $NR$  time instants and is constant for the remaining  $NU - NR$  instants

$$u_r(k) = \begin{cases} u_{i-1}(k) & \text{for } k = 1, \dots, NR \\ u_{i-1}(NR) & \text{for } k = NR + 1, \dots, NU \end{cases} \quad (3)$$

where the integer  $NR \leq NU$  is the length of the reference horizon. Thus,

$$U_r(i) = EU(i-1) \quad (4)$$

where

$$E = \begin{bmatrix} E_1 & 0 \\ E_2 & 0 \end{bmatrix}, \quad \underbrace{E_1 = \text{diag}(I_{nu}, \dots, I_{nu})}_{NR \text{ times}} \quad (5)$$

$$E_2 = \left. \begin{bmatrix} 0 & I_{nu} \\ \vdots & \vdots \\ 0 & I_{nu} \end{bmatrix} \right\} NR \text{ times}$$

and  $I_{nu}$  is the identity matrix of dimension  $nu$ .

The last input sequence, input-increment horizon  $\Delta U(i)$ , is defined with respect to reference horizon  $U_r(i)$  as follows:

$$\Delta U(i) = U(i) - U_r(i) = U(i) - EU(i-1) \quad (6)$$

Sequences  $U(i)$ ,  $U(i-1)$ ,  $U_r(i)$ , and  $\Delta U(i)$  are shown in Fig. 1.

Two output sequences are introduced: output horizon  $Y$

$$Y(i)^T = [y_i^T(1), y_i^T(2), \dots, y_i^T(NY)] \quad (7)$$

and predicted-output horizon  $\bar{Y}$

$$\bar{Y}^T(i) = [\bar{y}_i^T(1), \bar{y}_i^T(2), \dots, \bar{y}_i^T(NY)] \quad (8)$$

The latter is an output of the system with reference horizon  $U_r$  as an input. Components  $y_i(k)$  and  $\bar{y}_i(k)$  are the output and predicted output, respectively, at instant  $i$  with  $k$  steps ahead. Note that although the system output at instant  $i$  with  $k$  steps ahead is equal to the output at instant

$i + l$  with  $k - l$  steps ahead,  $y_i(k) = y_{i+1}(k - l) = y(i + l)$ ,  $l < k$ , the same is not true for the predicted output. The prediction at instant  $i$  with  $k$  steps ahead is not the same as the prediction at instant  $i + l$  with  $k - l$  steps ahead.

The output horizon is obtained from the plant model, Eq. (1), for  $k = 1, \dots, NY$ :

$$y(i + k) = CA^k x(i) + CA^{k-1} Bu(i) + \dots + CBu(i + k - 1) \quad (9)$$

Predicted output  $\bar{y}_i(k)$  is defined as a system response to the reference-horizon input  $U_r(i)$ . Thus, for  $k = 1, \dots, NY$

$$\bar{y}_i(k) = CA^k x(i) + CA^{k-1} Bu_r(1) + \dots + CBu_r(k) \quad (10)$$

and Eq. (9) is now

$$y(i + k) = \bar{y}(k) + CA^{k-1} B \Delta u_i(0) + \dots + (CA^{k-NR-1} B + \dots + CB) \Delta u_i(NR - 1) \quad (11)$$

Denoting the Markovian matrix

$$G = \begin{bmatrix} g_1 & 0 & \dots & 0 \\ g_2 & g_1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ g_{NU} & g_{NU-1} & \dots & g_1 \\ \dots & \dots & \dots & \dots \\ g_{NY} & g_{NY-1} & \dots & g_{NY-NU+1} \end{bmatrix} \quad (12)$$

where  $g_i = CA^{i-1} B$  is the  $i$ th Markov parameter, one obtains the output horizon from Eq. (11):

$$Y(i) = \bar{Y}(i) + G \Delta U(i) \quad (13)$$

Predicted-output horizon  $\bar{Y}$ , necessary to determine  $Y$  in Eq. (13), is determined from Eq. (10):

$$\begin{aligned} \bar{Y}(i) &= Hx(i) + GU_r(i) = Hx(i) + GEU(i - 1) \\ &= Hx(i) + FU(i - 1) \end{aligned} \quad (14)$$

where

$$F = GE, \quad H = \mathcal{O}_{NY-1} A \quad (15)$$

$E$  is as given by Eq. (5), and

$$\mathcal{O}_{NY-1} = [C^T (CA)^T \dots (CA^{NY-1})^T]^T \quad (16)$$

The input-increment horizon depends on the length of the reference horizon. In particular, for  $NR = 1$ , one obtains  $\Delta u_i(k) = u_i(k) - u_{i-1}(1)$ ,  $k = 1, \dots, NU$ . This is the case of the generalized predictive control of [5], where the control increments are defined with respect to the last input command. For  $NR = NU$ , one obtains  $E = I$  and  $\Delta U(i) = U(i) - U(i - 1)$ . In this case, the control increment is defined with respect to the previous control over the whole length of the input horizon. If the input increment is determined with respect to the zero-reference input, the input-increment horizon is equal to the input horizon,  $\Delta U(i) = U(i)$ , and for this case,  $NR = 0$ . Signal sequences  $U(i)$ ,  $U(i - 1)$ ,  $U_r(i)$ , and  $\Delta U(i)$  are shown in Fig. 1 for  $NR = 0, 1$ , and  $NU$  and for a generic  $NR$ . Note also that for  $NR = NU$  one obtains  $F = G$ , for  $NR = 1$  one obtains  $F^T = [g_1^T, g_1^T + g_1^T, \dots, g_1^T + \dots + g_{NY}^T]$ , and for  $NR = 0$  one obtains  $F = 0$ . In the latter case, the output is predicted from the system state only, while otherwise it is predicted from the state and the system input as well.

### III. Predictive Control

The basic task for a predictive controller is to assure that for the bounded input the future output  $Y$  will closely follow the input command  $Y_o$  within the output horizon  $NY$ :

$$Y_o^T(i) = [y_{oi}^T(1), y_{oi}^T(2), \dots, y_{oi}^T(NY)] \quad (17)$$

where  $y_{oi}(k)$  is the command signal at instant  $i$  with  $k$  steps ahead. Thus, the task is to minimize the plant tracking error while the input remains bounded:

$$\varepsilon(i) = Y_o(i) - Y(i) \quad (18)$$

$$\varepsilon^T(i) = [\varepsilon_i^T(1), \varepsilon_i^T(2), \dots, \varepsilon_i^T(NY)] \quad (19)$$

where  $\varepsilon_i(k)$  is the error at instant  $i$  with  $k$  steps ahead. The tracking error within the output horizon, as well as

the restrictions on the input within the input horizon, are included in performance index  $J$

$$J = \text{tr} (\varepsilon^T(i)Q\varepsilon(i) + \Delta U^T(i)R\Delta U(i)) \quad (20)$$

where  $\text{tr}(\cdot)$  denotes the trace of a square matrix;  $Q$  and  $R$  are symmetric, positive, definite matrices;  $Q$  is the tracking-error weighting matrix; and  $R$  is the control-effort weighting matrix.

The necessary condition for the optimum,  $\partial J/\partial \Delta U = 0$ , applied to Eq. (20) and using Eq. (13), yields

$$\Delta U(i) = K(Y_o(i) - \bar{Y}(i)) = K\bar{\varepsilon}(i) \quad (21)$$

where

$$K = (G^T Q G + R)^{-1} G^T Q \quad (22)$$

and  $\bar{\varepsilon}(i) = Y_o(i) - \bar{Y}(i)$  is the predicted output error. The resulting control increment  $\Delta U(i)$  covers the whole input-horizon  $NU$ ; for control purposes, however, only the first component (the current control increment) is used. Let  $k$  denote the first  $nu$  rows of  $K$ :

$$k = eK, \quad e = [I_{nu} \ 0 \ 0 \ \dots \ 0] \quad (23)$$

Then the control increment at instant  $i$  is

$$\Delta u(i) = k(Y_o(i) - \bar{Y}(i)) = k\bar{\varepsilon}(i) \quad (24)$$

and the control input is obtained from Eq. (6)

$$u(i) = u(i-1) + \Delta u(i) \quad (25)$$

Combining Eqs. (24), (25), (14), (15), and (16), one obtains the control command at moment  $i$ :

$$\begin{aligned} u(i) &= u(i-1) + k(Y_o(i) - \bar{Y}(i)) \\ &= u(i-1) + kY_o(i) - kHx(i) - kFU(i-1) \end{aligned}$$

and with  $u(i-1) = eU(i-1)$ , the above equation yields

$$u(i) = kY_o(i) - kHx(i) + (e - kF)U(i-1) \quad (26)$$

Thus, the command  $u(i)$  depends on the previous input horizon  $U(i-1)$ , on the actual state  $x(i)$ , and on the control command  $Y_o(i)$  up to  $NY$  steps ahead.

The closed-loop system equations are obtained by combining the plant equation, Eq. (1), with the controller equation, Eq. (26), and by introducing the new state variable  $U_o(i)$  such that

$$U_o(i+1) = U(i) \quad (27)$$

In this way, one obtains

$$x(i+1) = (A - BkH)x(i) + B(e - kF)U_o(i) + BkY_o(i)$$

$$U_o(i+1) = -KHx(i) + (I_N - KF)U_o(i) + KY_o(i)$$

$$y(i) = Cx(i) \quad (28)$$

and  $N = NU \times nu$ . With the new state variable  $z^T = [x^T, U_o^T]$ , the closed-loop equations are

$$z(i+1) = A_c z(i) + B_c Y_o(i), \quad y(i) = C_c z(i) \quad (29)$$

where

$$\begin{aligned} A_c &= \begin{bmatrix} A - BkH & B(e - kF) \\ -KH & I_N - KF \end{bmatrix} \\ &= \begin{bmatrix} (I_n - BK \mathcal{O}_{NY-1})A & B(e - kF) \\ -KH & I_N - KF \end{bmatrix} \end{aligned} \quad (30)$$

$$B_c^T = [Bk \ K]^T, \quad C_c = [C \ 0] \quad (31)$$

One can see that control command  $u(i)$  is now

$$u(i) = kY_o(i) + [-kH \ e - kF]z(i) \quad (32)$$

fully recovered from the current state of the system and from the input command.

The block diagram for the closed-loop system, Eqs. (29)–(31), is presented in Fig. 2. The system consists of the plant, the predictor (PRD), the controller (CO),

and the command horizon generator (CHG). The predictor structure is shown in Fig. 3(a) and the command horizon generator in Fig. 3(b). The state-space representation of the command horizon generator is

$$A_{rh} = \begin{bmatrix} 0 & I & 0 & \dots & 0 \\ 0 & 0 & I & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & I \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad B_{rh} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \end{bmatrix}, \quad (33)$$

$$C_{rh} = \begin{bmatrix} I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & I \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad D_{rh} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \end{bmatrix}$$

where  $I$  is an identity matrix of order  $ny$ , and  $A_{rh}$  and  $B_{rh}$  have  $NY - 1$  rows, while  $C_{rh}$  and  $D_{rh}$  have  $NY$  columns.

The weighting matrices  $R$  and  $Q$  are the tuning parameters of the optimal design (for example, see [2,12]). That means they are to be adjusted until satisfactory results are obtained. Although they are not "active" in the optimal design solution, their choice significantly influences the performance and stability of the system. A general procedure for a reasonable choice of the weighting matrices is not yet known. In this article, a simplified procedure is developed. The weighting matrices obtained from this procedure significantly improve system performance, i.e., tracking error.

A diagonal matrix  $R = \rho I$  has been chosen as a control weighting matrix, where  $\rho > 0$  is a scalar, and tracking-error weighting matrix  $Q$  has the following structure:

$$Q = \text{diag}(q, \alpha q, \alpha^2 q, \dots, \alpha^{NY-1} q) \quad (34)$$

The diagonal component  $q\alpha^{k-1}$  is the weight of the error of  $\varepsilon_i(k)$ , the  $k$ th component of  $\varepsilon(i)$ . The last weight is time dependent; the weight of the output error at the  $(i+k)$ th time instant is  $\alpha^{k-1}$ . The scalar  $\alpha$  is called a forgetting factor. The most recent output is given a unit weight, and the future output penalized (in fact, awarded, as will be shown later) exponentially. With this arrangement, the choice of  $R$  and  $Q$  reduces to the choice of parameters  $\rho$ ,  $\alpha$ , and  $q$ , as is illustrated in Section V.B.

There are two sources of system disturbances: measurement noise  $v_y(i)$  (or  $v_x(i)$  when measuring all state

variables) and input disturbances  $v_u(i)$  (Fig. 2). The disturbances are included in the closed-loop system model, with the triple  $(A_c, B_y, C_c)$  for the output noise and the triple  $(A_c, B_u, C_c)$  for the input disturbances, where

$$B_y = \begin{bmatrix} A - BkH \\ -KH \end{bmatrix},$$

$$B_u = \begin{bmatrix} B(e - kF)e_v \\ (I - KF)e_v \end{bmatrix}, \quad e_v = \begin{bmatrix} I_{nu} \\ 0 \end{bmatrix}$$

Their impact on system performance is studied in Section V.D. For high-frequency disturbances, the disturbance-rejection properties of the system significantly improve when the lowpass filter (LPF) is applied as in Fig. 4(a). The plant states related to the command signal are obtained from the plant model (PM), and they are extracted from the measured states. The resulting signal passes through a lowpass filter and is added to the states previously extracted. The filter is shown in Fig. 4(b).

#### IV. Predictive Estimation

Implementation of the predictive controller depends on the availability of the plant states for measurement. Often, these parameters are not available. An LQ estimator (for example, see [2,13]) that estimates plant state from its output can be considered as a solution to the problem. Its action, however, is too slow for the predictive control system, and the predictive scheme is included in the design of the estimator. Thus, a new estimator with dynamic characteristics comparable to the predictive controller is developed.

The estimate  $\hat{x}(i)$  of the plant state  $x(i)$  is determined from the input and output horizons as follows. From Eq. (1), one obtains

$$CA^k x(i) = y(i+k) - \sum_{j=1}^k CA^{j-1} Bu(i+k-j),$$

$$k = 0, 1, 2, \dots, NY$$

or

$$\mathcal{O}_{NY} x(i) = Y_e(i) - G_e U(i) \quad (35)$$

where

$$\mathcal{O}_{NY} = \begin{bmatrix} C \\ H \end{bmatrix}, \quad Y_e(i) = \begin{bmatrix} y(i) \\ Y(i) \end{bmatrix}, \quad G_e = \begin{bmatrix} 0 \\ G \end{bmatrix} \quad (36)$$

and  $H$ ,  $G$ , and  $Y$  are given in Eqs. (15), (12), and (7), respectively. Variable  $Y_e$  is an augmented output horizon composed of the current output  $y(i)$  and the output horizon  $Y(i)$ . From Eq. (35), the estimate  $\hat{x}(i)$  of  $x(i)$  is determined such that for a symmetric, positive, weighting matrix  $Q_e$  the estimation index

$$J_e = \|H_e x(i) - H_e \hat{x}(i)\|_{Q_e}^2 \quad (37)$$

is minimal, obtaining

$$\hat{x}(i) = \mathcal{O}_{NY}^+ (Y_e(i) - G_e U(i)) \quad (38)$$

where  $\mathcal{O}_{NY}^+ = (\mathcal{O}_{NY}^T Q_e \mathcal{O}_{NY})^{-1} \mathcal{O}_{NY}^T Q_e$ . Note that the state estimate is determined from input and output horizons, while input and output signals, rather than horizons, are available for estimation. The input horizon is available, however, right after the controller output (Fig. 2). Output horizon  $Y(i)$  is not available directly; nevertheless, it can be obtained from the plant model as follows:

$$X(i+1) = AX(i) + BU(i), \quad Y(i) = CX(i) + DU(i) \quad (39)$$

The estimator is shown in Fig. 5(a). Thus, the plant state is estimated from its output and the input horizon. This scheme is similar to the LQ estimation scheme, since it uses the available input and output signals and the plant model to generate the estimate. The block diagram of the predictive control system with the predictive estimator (EST) is shown in Fig. 6.

Unlike the LQ estimator, the predictive estimator does not have filtering properties, since its output  $\hat{x}(i)$  is proportional to a noisy signal  $y(i)$ . This drawback can be removed as follows. Given the plant model output  $y_n(i)$  the output error  $\varepsilon_y(i) = y(i) - y_n(i)$  is filtered by a proper filter, obtaining the filtered error  $\varepsilon_{yf}(i)$ . In most cases, the output error is a high-frequency noise; hence, a lowpass filter is applied. The filtered output is obtained by adding a filtered error to the nominal output  $y_f(i) = y_n(i) + \varepsilon_{yf}(i)$ . In this way, most of the noise power is removed from the output signal, while the basic properties of the signal remain untouched. The estimator with a filter is shown in Fig. 5(b). The filter action will be illustrated in the next section.

## V. Predictive Control and Estimation for the DSS 14 Antenna

Performances of the predictive controller and estimator are checked through tracking simulations of the DSS 14 National Aeronautics and Space Administration (NASA)/Jet Propulsion Laboratory (JPL) 70-meter DSN antennas. The existing control scheme for the DSN 70-m antennas [14] is based on an LQ regulator design with the integral action as presented in [15–20]. The LQ control system is shown in Fig. 7, in which the plant output is augmented by the addition of the output integrals in order to ensure the zero mean value of the constant-rate tracking error. The LQ controller is designed for this augmented plant with a constant tracking command. This assumption can be a significant source of tracking error. A controller designed for the constant tracking command can result in insufficient antenna performance, especially for relatively fast commands or varying rate commands. In this section, the performance of the predictive controller is compared with that of the LQ controller in the tracking environment.

### A. Plant Model

The state-space model of the DSS 14 antenna [14] is a four-state model with position rate  $u$  as an input and position rate  $y$  as an output. Its discrete-time representation  $(A_d, B_d, C_d)$ , with the sampling period  $\Delta t = 0.05$  sec, is obtained from the continuous-time representation in [2]:

$$A_d = \begin{bmatrix} 0.0468 & 0 & 0 & 0 \\ 0 & 0.5443 & 0.3474 & 0 \\ 0 & -0.3474 & 0.5443 & 0 \\ 0 & 0 & 0 & 0.8872 \end{bmatrix},$$

$$B_d = \begin{bmatrix} 0.0113 \\ 0.0025 \\ 0.0399 \\ 0.0538 \end{bmatrix}, \quad C_d = [0.7239 \quad 9.2260 \quad 0 \quad 1.1421]$$

The system  $(A_d, B_d, C_d)$  is augmented. As a result, its output consists of the position rate, the angular position, and the integral of the position. The augmented system is shown in Fig. 8. Denoting  $x_d$  the state of the system  $(A_d, B_d, C_d)$ , and  $x_{po}$  and  $x_{ipo}$  the position and the integral of the position, respectively, one obtains from Fig. 8

$$x_{ipo}(i+1) = x_{ipo}(i) + \Delta t x_{po}(i)$$

$$x_{po}(i+1) = x_{po}(i) + \Delta t C_d x_d(i)$$

$$x_d(i+1) = A_d x_d(i) + B_d u(i)$$

With the plant-state variable  $x^T = [x_{ip0}, x_{po}, x_d^T]$ , the triple  $(A, B, C)$  is the resulting plant-state space representation

$$A = \begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & \Delta t C_d \\ 0 & 0 & A_d \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ B_d \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & C_d \end{bmatrix}$$

used in simulations presented below.

## B. Weighting Matrices and Input and Output Horizons

For simulation purposes, a piecewise-linear profile of the position command is chosen, with linear increase followed by linear decrease and the final constant value (Fig. 9). The command rate is 4 mdeg/sec, which is a typical sidereal tracking rate. The shape of the command is more dramatic than the real tracking command, but it has been chosen to emphasize the tracking possibilities of the predictive controller. A more realistic tracking command will be used later in this article.

In order to perform a series of simulations, weighting matrices  $R$  and  $Q$  are chosen such that the output error is small while the control effort is maintained within reasonable limits. For diagonal-control weighting matrix  $R = \rho I$ , parameter  $\rho = 0.01$  is chosen. Tracking-error weighting matrix  $Q$  is as in Eq. (34). Component  $q$  is in the form  $q = \text{diag}(q_i, q_p, q_r)$ . It represents the weight of the integral, position, and rate components of the output. The following choices of weight are recommended from a series of simulations tracking the command as in Fig. 9: for the integral-of-the-position signal,  $q_i = 10$ ; for the position signal,  $q_p = 1$ ; and for the rate signal,  $q_r = 0.1$ . Coefficient  $\alpha^{k-1}$  in the weighting matrix is the weight of the  $k$ th error component in the output horizon. Simulations have been performed in order to determine the value of parameter  $\alpha$ . The plot of the Euclidean norm of tracking error  $\|y - y_o\|_2$  versus  $\alpha$  is shown in Fig. 10 both for different lengths of output horizon and for lengths of input and reference horizon equal to lengths of output horizon. The plot shows the minimal tracking error obtained for  $\alpha = 6.2$  and  $NY = NU = NR = 6$ . For  $NR = NU = NY$ , Fig. 11 plots the values of  $\alpha$  for which the tracking error is minimal. The figure shows that, for horizons that are long enough, the forgetting factor is close to 1. Thus, in this case, the time weighting does not significantly improve the tracking error. However, for short horizons, the proper choice of forgetting factor is a critical factor that minimizes the error dramatically. In prediction, the forgetting factor is greater than one (thus, "reminding factor" could be an adequate name for it). This is in contrast to the forgetting factor value in estimation procedures such as in

[12], where the factor is smaller than one. This difference occurs due to opposite time directions; in estimation, the past values of signal are weighted, while in prediction the future values of signal are processed.

From simulations, the impact of the length of input-reference horizon  $NR$  and output horizon  $NY$  on the tracking error is determined. The results are plotted in Fig. 12. One can see that for  $NR \geq n/2$  and  $NY \geq 2n$  (where  $n = 6$  is the number of plant-state variables), the performance error is close to the minimal one.

## C. Antenna Performance

The performance of the DSS 14 antenna with the tracking command as shown in Fig. 9 has been evaluated for the parameters recommended above. The following parameters for the predictive controller were chosen:  $NR = NU = NY = n = 6$  and weighting matrices with  $\rho = 0.01$ ,  $q = \text{diag}(10, 1, 0.1)$ , and  $\alpha = 6.2$ . The reference signal and the position of the antenna with the predictive controller for  $\alpha = 6.2$ , for  $\alpha = 1$ , and for the antenna with the LQ controller are shown in Fig. 9. The prediction errors and control input for the above three cases ( $\alpha = 6.2$ ,  $\alpha = 1$ , and LQ controller) are shown in Figs. 13 and 14. The figures show better performance by the predictive controllers than by the proposed LQ controller with comparable control effort. Also, predictive-controller performance with time-weighted output error ( $\alpha > 1$ ) is better than predictive-controller performance without time weighting ( $\alpha = 1$ ).

The minimum of tracking error for output horizon  $NY = 4$  is obtained for  $\alpha = 6.4$ . These two parameters are used in further simulations, since it is reasonable to have the length of the output horizon as small as possible; the dimension of the controller as well as the complexity of the system depend on  $NY$ . The step-response and frequency-response plots of the closed-loop system with a predictive controller and an LQ controller are compared in Figs. 15 and 16. Figure 15 shows that the settling time and overshoot for the system with the predictive controller, with  $NY = NU = NR = 4$  and  $\alpha = 6.4$ , are significantly reduced from the system with the LQ controller. Similarly, from Fig. 16, one can see the tracking performance is improved; the magnitude of the closed-loop transfer function is equal to 1 over a wider bandwidth. Also, roll-off rate is improved for the system with a predictive controller (with  $NY = NU = NR = 4$  and  $\alpha = 6.4$ , as well as with  $NY = NU = NR = 6$  and  $\alpha = 5$ ) when compared to the system with an LQ controller.

The piecewise constant-rate command, as well as the unit step command, are dramatic scenarios for the DSS 14

antenna and have been introduced in order to present dynamic possibilities of the predictive control. In order to meet the typical working requirements for the antenna, the raised-cosine command is introduced, as in Fig. 17 (solid line). This kind of command is close to the real elevation or azimuth trajectory of the antenna (conscan-like tracking). The plot of the output of the predictive control system overlaps the plot of the command, while the output of the LQ control system is plotted by a dashed line (Fig. 17). The tracking error, the difference between the output and the command, is plotted in Fig. 18(a) for the LQ control system and in Fig. 18(b) for the predictive control system. For the LQ control system, the error is on the order of  $10^{-4}$ , while the error for the predictive control system is on the order of  $10^{-7}$ . In both cases, however, the control effort is almost the same (Fig. 19).

#### D. Robustness and Disturbance Suppression

The robustness of the closed-loop system to the plant-parameter variations is checked as follows. The plant poles are randomly perturbed within 20 percent margin, and the error in the step-command tracking is simulated for 500 random samples. The results of the simulations are presented in Fig. 20. On the average, tracking error has changed about 5 percent in comparison with the nominal plant error, and the maximal tracking error is 66 percent larger than the nominal plant error. The step and frequency responses of the closed-loop system for the nominal plant and for the plant model deviated 20 percent from the nominal, as is shown in Figs. 21 and 22. Both plots show good performance and robustness of the system.

Two sources of disturbances of the antenna are studied: the input disturbances and output disturbances (measurement noise). The input-disturbance transfer functions (from  $v_u$  to  $y$ ) and output-disturbance transfer functions (from  $v_x$  to  $y$ ) are shown in Fig. 23, the latter one for both position- and rate-measurement noise. One can see from Fig. 23 that the input disturbances are significantly suppressed, while the position-measurement noise is amplified over certain frequency ranges.

The nature of the antenna disturbances is not satisfactorily known, and here their general properties are outlined. Input disturbances, such as wind or thermal forces, are low-frequency signals. Measurement noise, on the other hand, is a high-frequency signal (high in comparison to the antenna fundamental frequency, which is less than 1 Hz). Therefore, for testing purposes, white-noise input disturbances and high-frequency measurement noise with frequency components over 3 Hz are applied. The system response due to different signal-to-noise (S/N) ratios

is simulated. The results are compared in Fig. 24, where good disturbance-suppression properties of the system are observed. The impact of the input noise is much smaller than that of measurement noise. This feature can be explained with the lowpass-filtering property of the plant; the noise is filtered before entering the predictor. The tracking error and plant input for input noise with S/N ratio = 10 and for measurement noise with S/N ratio = 100 are shown in Figs. 25 and 26. The effect of the measurement noise is reduced by applying a filter, as in Fig. 4. The transfer-function plots from the output disturbances to the system output for the system with the filter are shown in Fig. 27. The tracking error due to measurement noise is reduced significantly, even for white noise (Fig. 28).

#### E. Predictive Estimator

Predictive-estimator performance is compared to the performance of the LQ estimator. The plant model ( $A_d, B_d, C_d$ ) has been used for simulations, with unit-step input and zero-initial conditions. For estimation purposes, the initial conditions have been changed to  $[0.1 \ 0.1 \ 0.1 \ 0.1]^T$ . The estimation results are shown in Fig. 29. The LQ estimator needs approximately 2 seconds to reach an acceptable estimation error, while the predictive estimator determines the states in virtually no time. In the case of noisy output, with S/N ratio = 100, one obtains estimation errors for the LQ and predictive estimators (with and without filter) as in Fig. 30. The unfiltered predictive estimate in Fig. 30(b) would make the estimator useless for prediction purposes. However, these errors are reduced by a filter, as in Fig. 30(c). The maximum error of the predictive estimator with a filter is much smaller than the residual error of the LQ estimator even after 4 seconds in action. Finally, simulations indicate that the identity-weighting matrix ( $Q_e = I$ ) is the optimal choice for estimation purposes.

### VI. Conclusions

In this article, a modified state-space predictive controller is introduced, and a predictive estimator is presented to complement the design of a predictive-control law. This approach has been used for the design of the tracking controllers of the NASA/JPL 70-m antennas. Several tracking scenarios have been introduced (step input, constant-rate rise and fall, raised-cosine trajectory) to test the tracking behavior of the predictive controller. Significant improvement of performance for presented scenarios has been observed. It has been shown that the time for the predictive estimator to reach an acceptable level of estimation error is much smaller than that for the LQ estimator. Also, a wider bandwidth and improved roll-off

rate is obtained for the predictive closed-loop system in comparison with the LQ regulator system. The predictive control system is robust to the plant-parameter variations. Shifts of plant poles of 20 percent of their nominal values keep the tracking performance good; the tracking error is on the same order as for a nominal plant. Disturbance-

suppression properties of a predictive control system also have been simulated and found to be good for input disturbances and measurement noise if the measurement-noise spectrum is higher than the plant-fundamental frequency. The system disturbance-suppression properties can be enhanced if the disturbance filter is included in the system.

## Acknowledgments

The author wishes to thank Ben Parvin for recognizing the importance of predictive control in DSN antenna design and for introducing technical aspects of DSN antenna controls, and Leon Alvarez for discussions about the DSS 14 antenna model.

## References

- [1] P. Albertos and R. Ortega, "On Generalized Predictive Control: Two Alternative Formulations," *Automatica*, vol. 25, no. 5, pp. 753-755, October 1989.
- [2] B. D. O. Anderson and J. B. Moore, *Optimal Control*, Englewood Cliffs, New Jersey: Prentice-Hall, 1990.
- [3] K. J. Astrom and B. Wittenmark, *Adaptive Control*, Reading, Pennsylvania: Addison-Wesley, 1989.
- [4] D. W. Clarke and C. Mohtadi, "Properties of Generalized Predictive Control," *Automatica*, vol. 25, no. 6, pp. 859-875, December 1989.
- [5] D. W. Clarke, C. Mohtadi, and P. S. Tuffs, "Generalized Predictive Control, Part I and II," *Automatica*, vol. 23, no. 2, pp. 137-160, April 1987.
- [6] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*, Englewood Cliffs, New Jersey: Prentice-Hall, 1984.
- [7] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, New York: Wiley, 1972.
- [8] F. L. Lewis, *Optimal Control*, New York: Wiley-Interscience, 1986.
- [9] R. Ortega and G. S. Galindo, "Globally Convergent Multistep Receding Horizon Adaptive Controller," *Int. J. Control*, vol. 49, no. 5, pp. 1655-1664, May 1989.
- [10] J. G. Reid, D. E. Chaffin, and J. T. Silverthorn, "Output Predictive Algorithmic Control: Precision Tracking with Application to Terrain Following," *J. Guidance, Control, and Dynamics*, vol. 4, no. 5, pp. 502-509, October 1981.
- [11] Y. Xi, "New Design Method for Discrete-Time Multi-Variable Predictive Controllers," *Int. J. Control*, vol. 49, no. 1, pp. 45-56, January 1989.
- [12] J. M. Maciejowski, *Multivariable Feedback Design*, Wokingham, England: Addison-Wesley, 1989.

- [13] F. L. Lewis, *Optimal Estimation: With an Introduction to Stochastic Control Theory*, New York: Wiley-Interscience, 1986.
- [14] L. S. Alvarez and J. Nickerson, "Application of Optimal Control Theory to the Design of the NASA/JPL 70-Meter Antenna Axis Servos," *TDA Progress Report 42-97*, vol. January-March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 112-126, May 15, 1989.
- [15] M. Athans, "On the Design of PID Controllers Using Optimal Linear Regulator Theory," *Automatica*, vol. 7, no. 6, pp. 643-647, December 1971.
- [16] S. Fukata, A. Mohri, and M. Takata, "On the Determination of the Optimal Feedback Gains for Multivariable Linear Systems Incorporating Integral Action," *Int. J. Control*, vol. 31, no. 6, pp. 1027-1040, June 1980.
- [17] C. D. Johnson, "Optimal Control of the Linear Regulator with Constant Disturbances," *IEEE Trans. Automatic Control*, vol. 13, no. 4, pp. 416-421, August 1968.
- [18] B. Porter, "Optimal Control of Multivariable Linear Systems Incorporating Integral Feedback," *Electronics Letters*, vol. 7, no. 8, pp. 170-172, April 1971.
- [19] T. Yahagi, "Optimal Output Feedback Control in the Presence of Step Disturbances," *Int. J. Control*, vol. 26, no. 5, pp. 753-762, November 1977.
- [20] P. C. Young and J. C. Willems, "An Approach to the Linear Multivariable Servomechanism Problem," *Int. J. Control*, vol. 15, no. 5, pp. 961-979, February 1972.

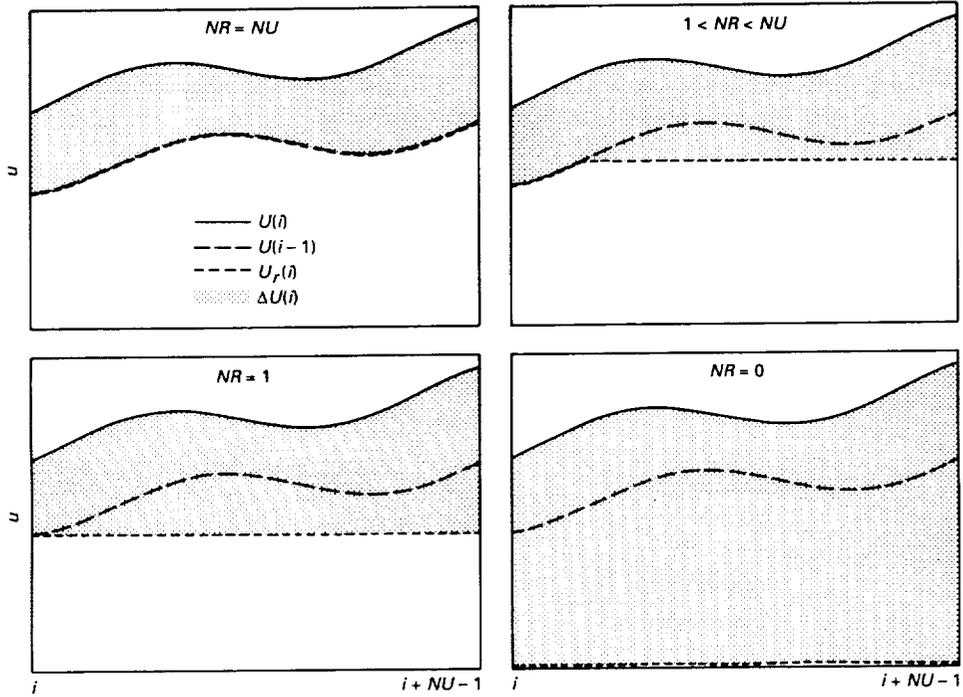


Fig. 1. Input horizons.

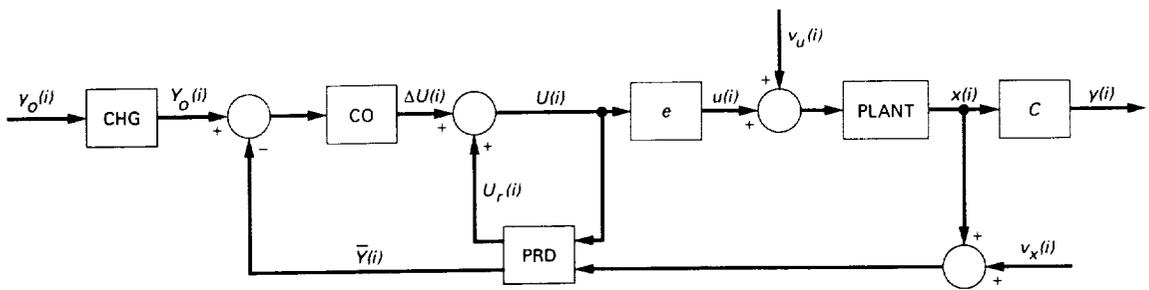


Fig. 2. The predictive control system.

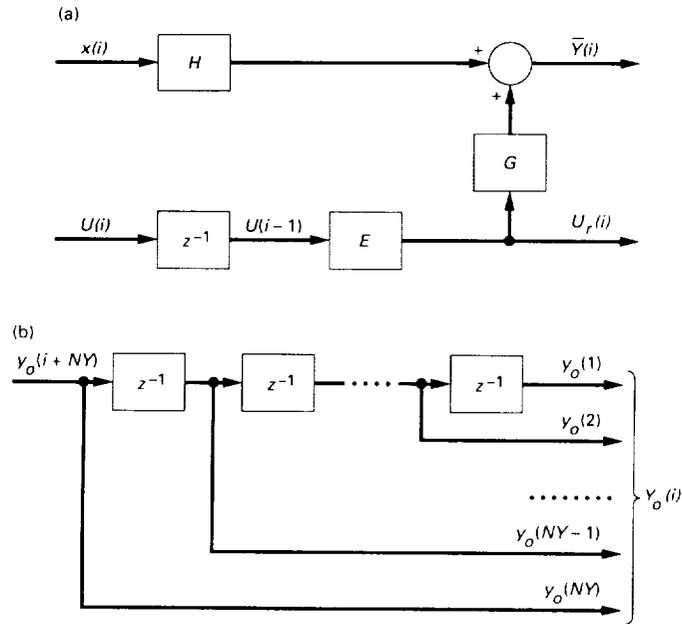


Fig. 3. Block diagrams of (a) the predictor, and (b) the command horizon generator.

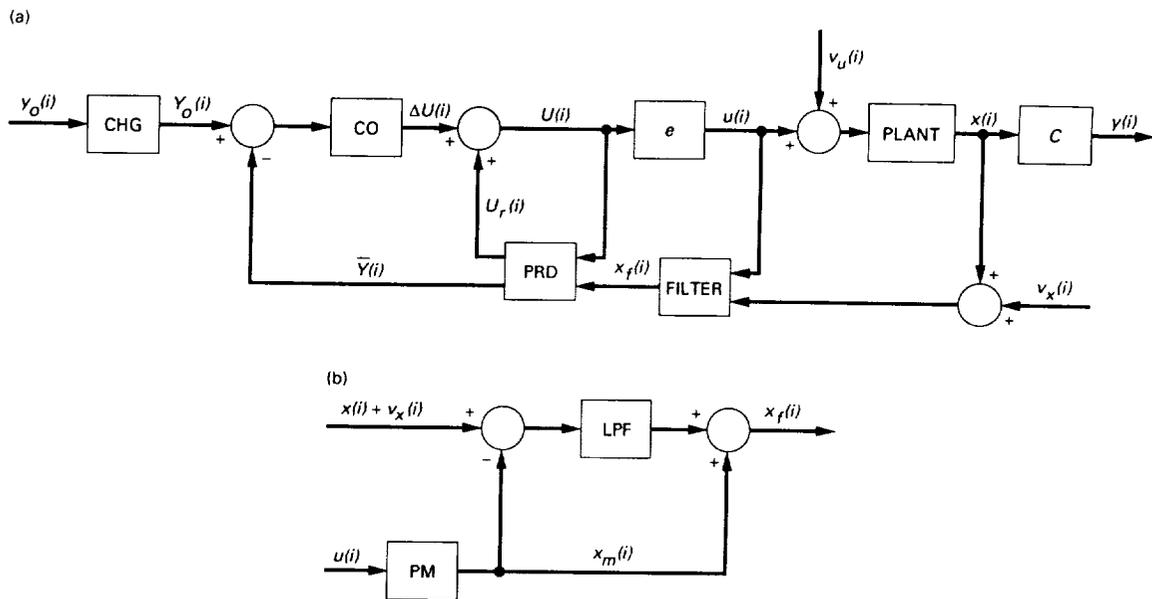


Fig. 4. Block diagrams of (a) the predictive control system with a measurement noise filter, and (b) the measurement noise filter.

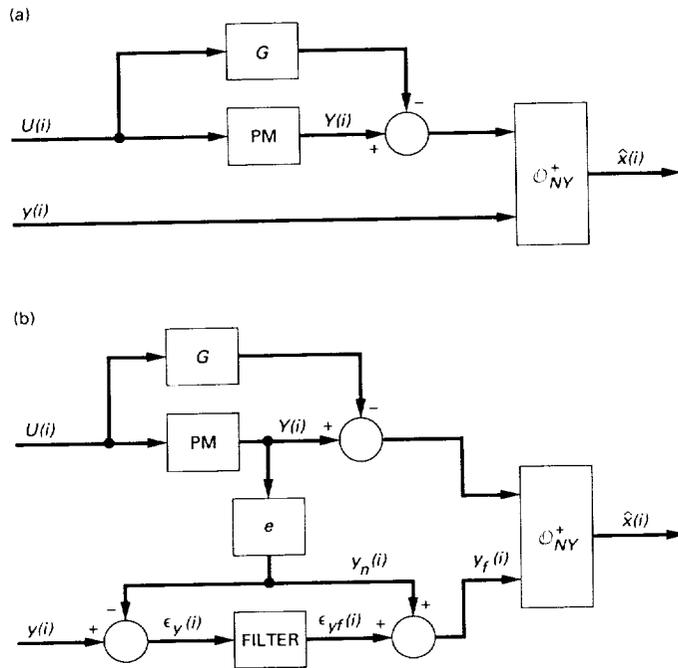


Fig. 5. The predictive estimator: (a) without measurement noise filter, and (b) with measurement noise filter.

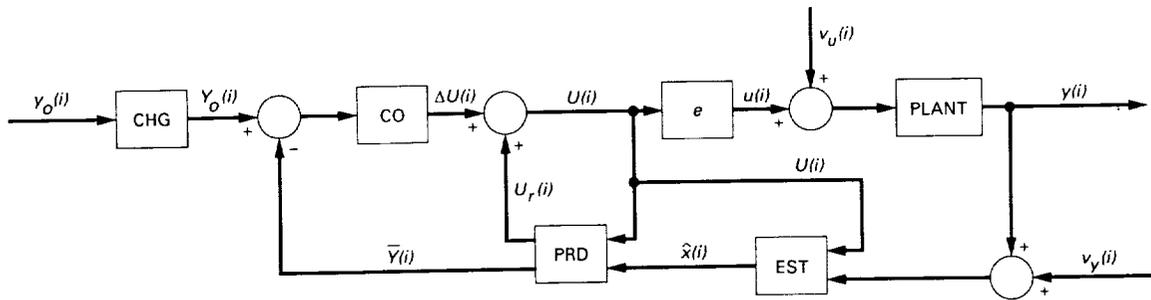


Fig. 6. The predictive control and estimation system.

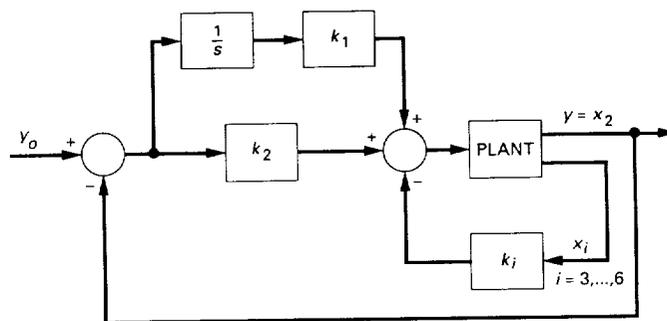


Fig. 7. The LQ control system for the DSS 14 antenna.

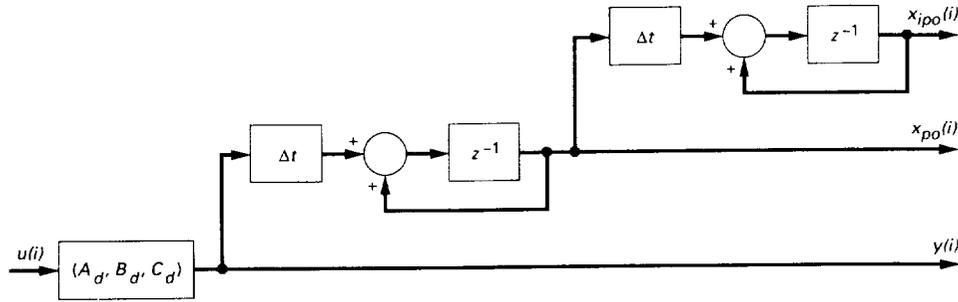


Fig. 8. The augmented model of the DSS 14 antenna.

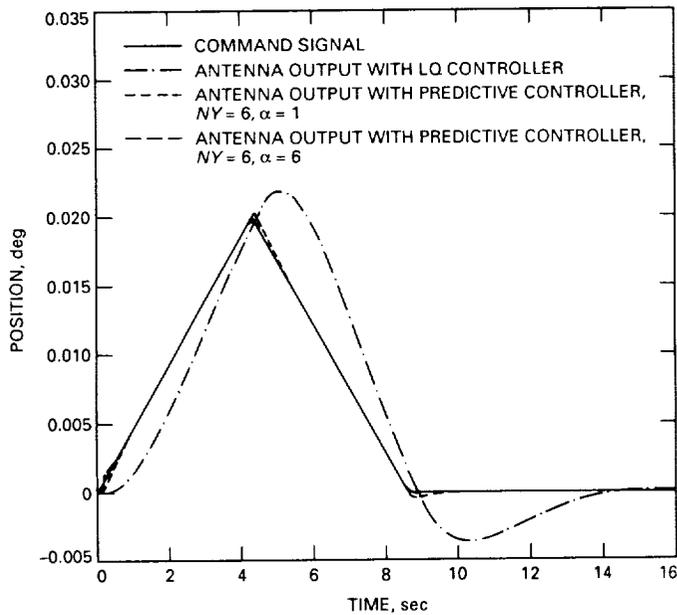


Fig. 9. The command signal and antenna output.

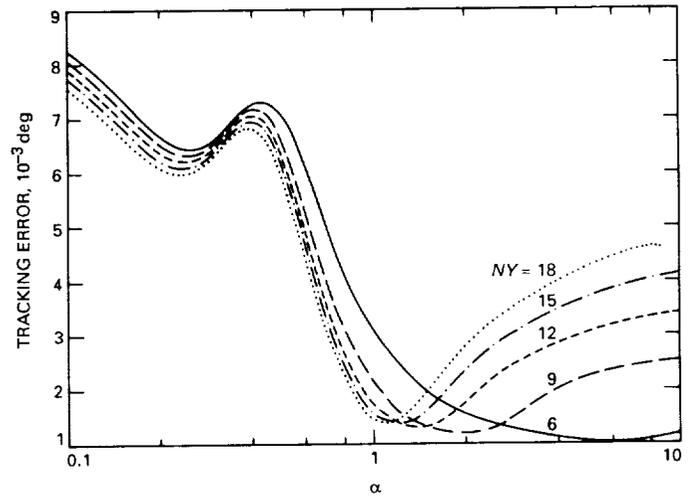


Fig. 10. Tracking error versus forgetting factor  $\alpha$  for different lengths of output horizon  $NY$ .

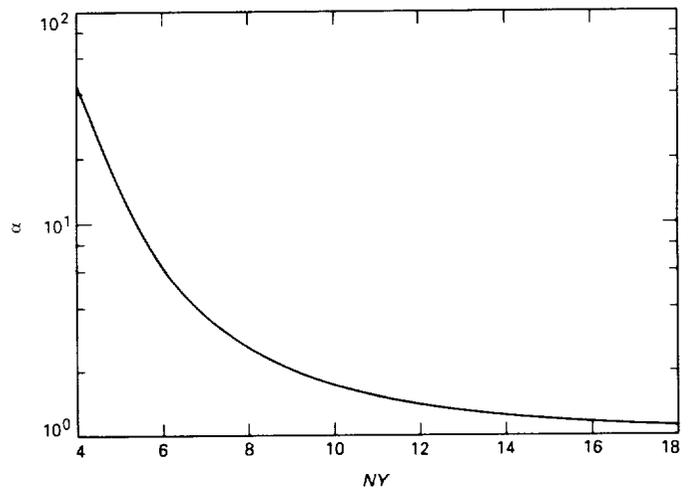


Fig. 11. Forgetting factor  $\alpha$ , for which the minimal tracking error is achieved versus length of output horizon  $NY$ .

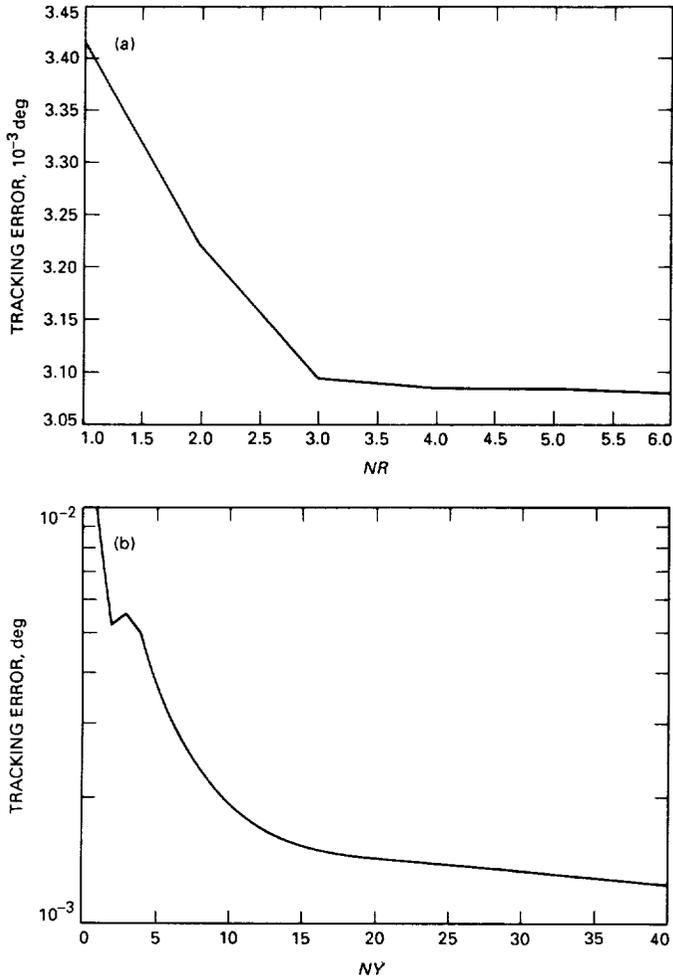


Fig. 12. Tracking error versus: (a) length of reference horizon  $NR$ , and (b) length of output horizon  $NY$ .

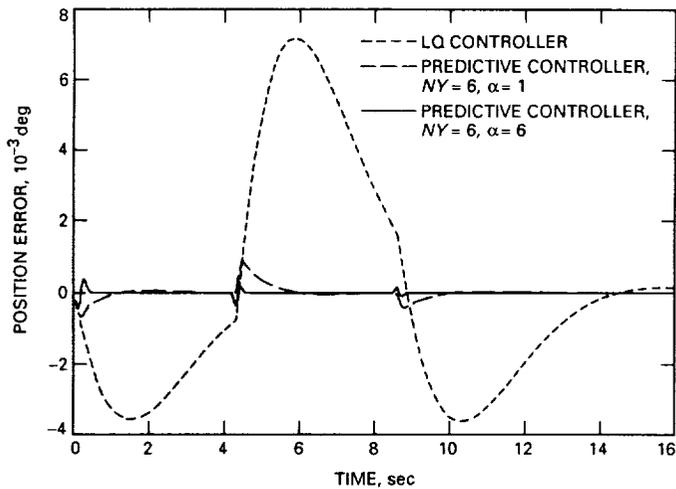


Fig. 13. Tracking error.

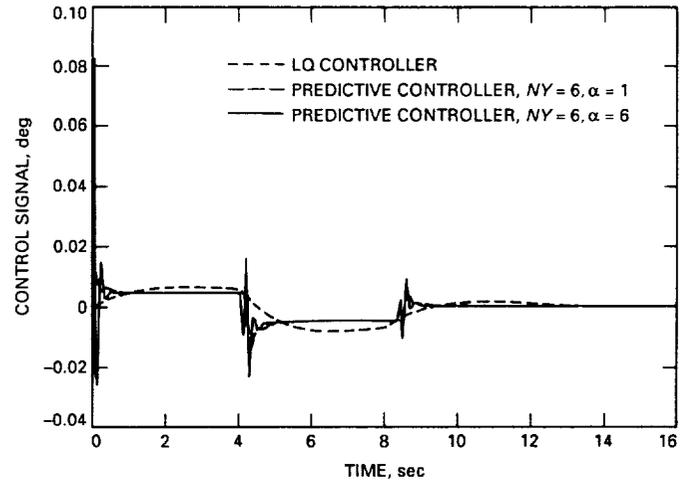


Fig. 14. Antenna control input.

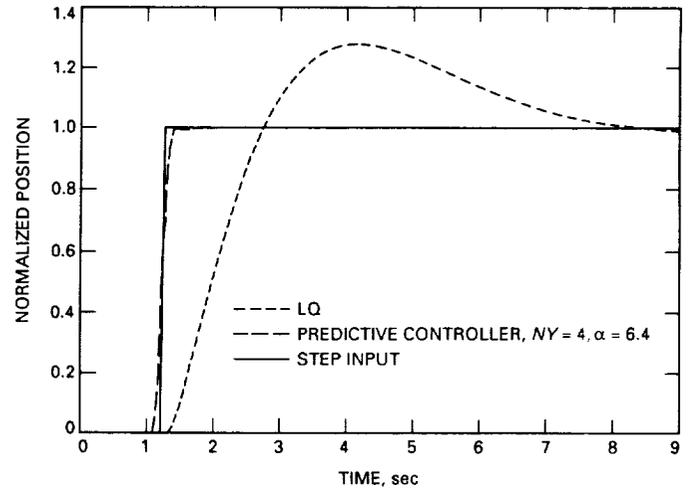


Fig. 15. Closed-loop system step response.

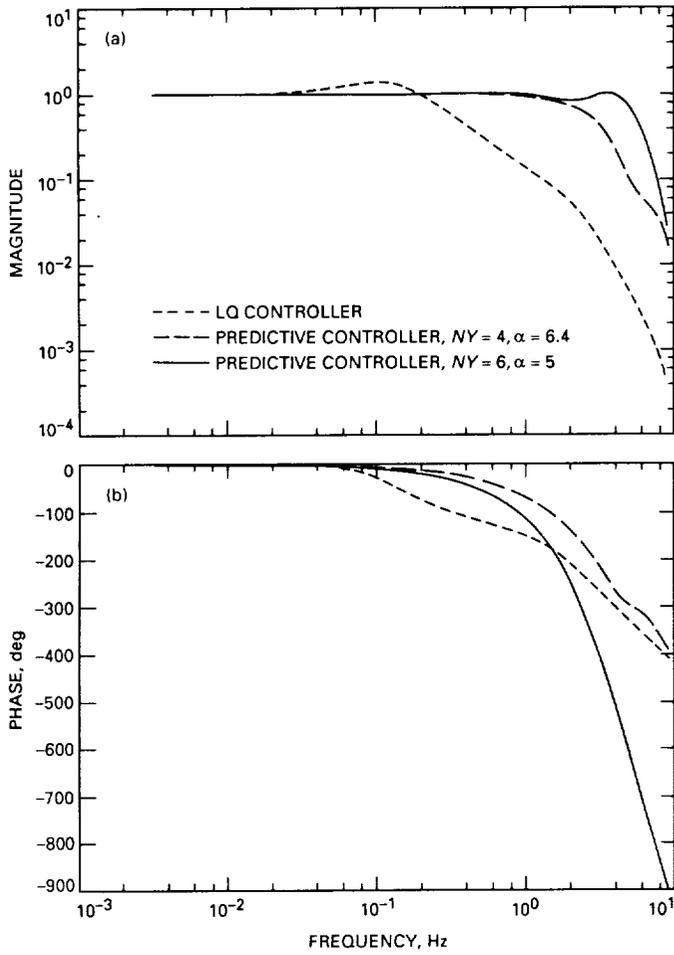


Fig. 16. Closed-loop system frequency response: (a) magnitude, and (b) phase.

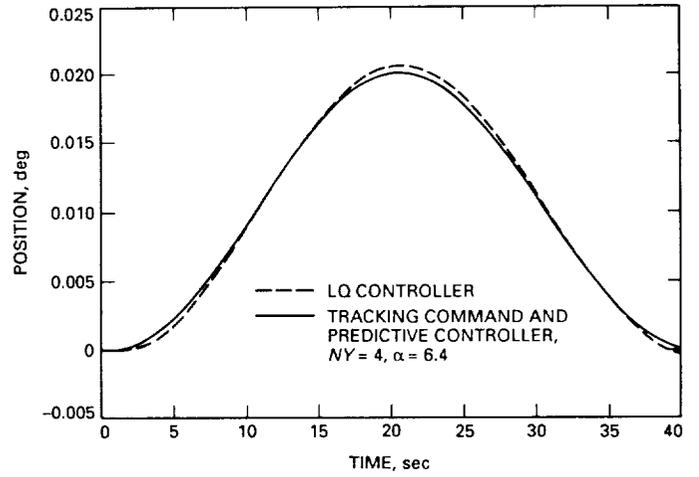


Fig. 17. Closed-loop system tracking performance.

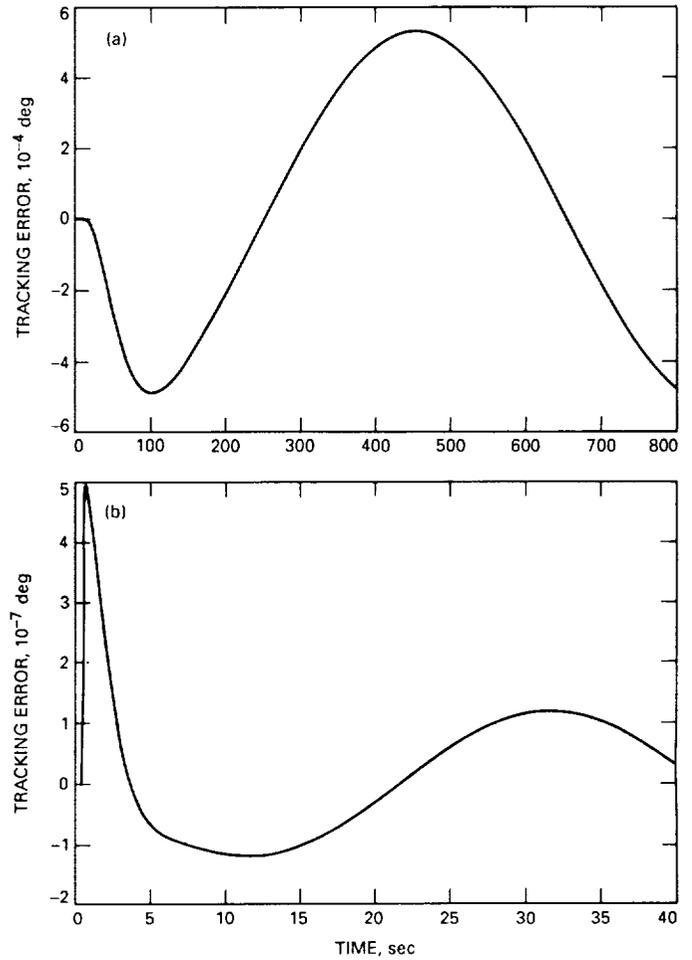


Fig. 18. Tracking error: (a) the LQ regulator system, and (b) the predictive system with  $NY = 4, \alpha = 6.4$ .

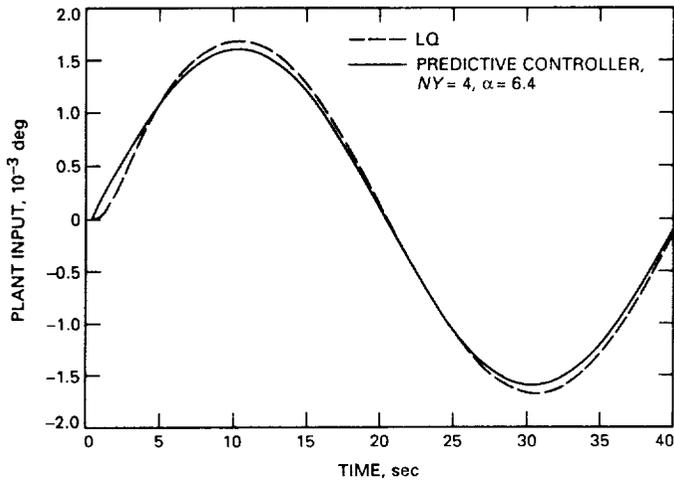


Fig. 19. Antenna input.

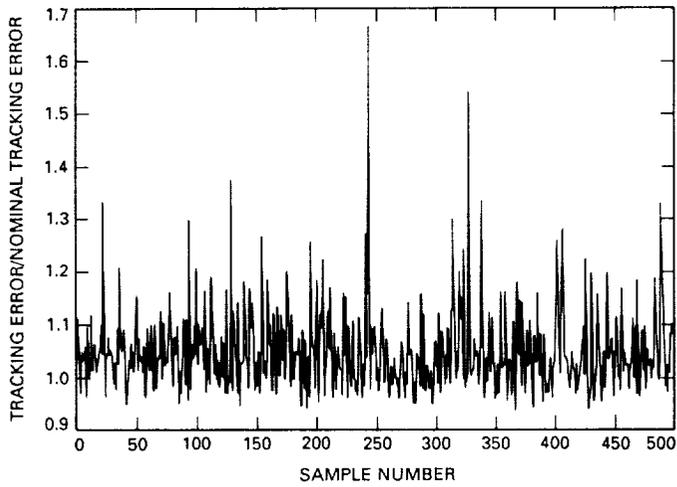


Fig. 20. Tracking-error ratio of the 20-percent randomly deviated model to the nominal model.

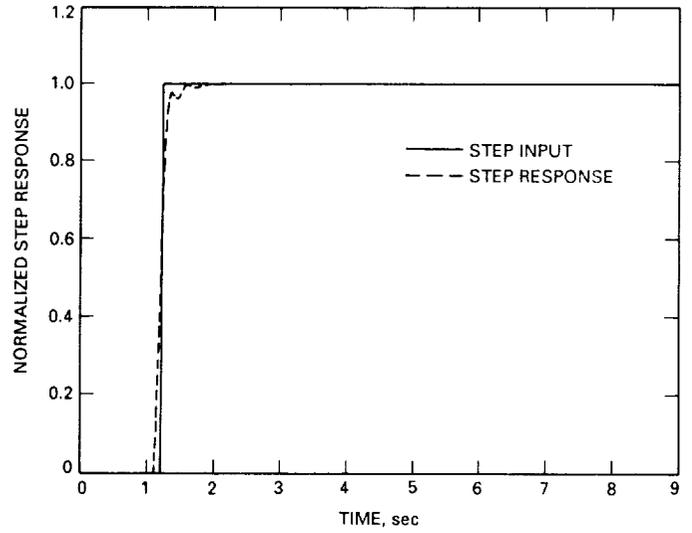


Fig. 21. Closed-loop step response with the 20-percent-deviated model.

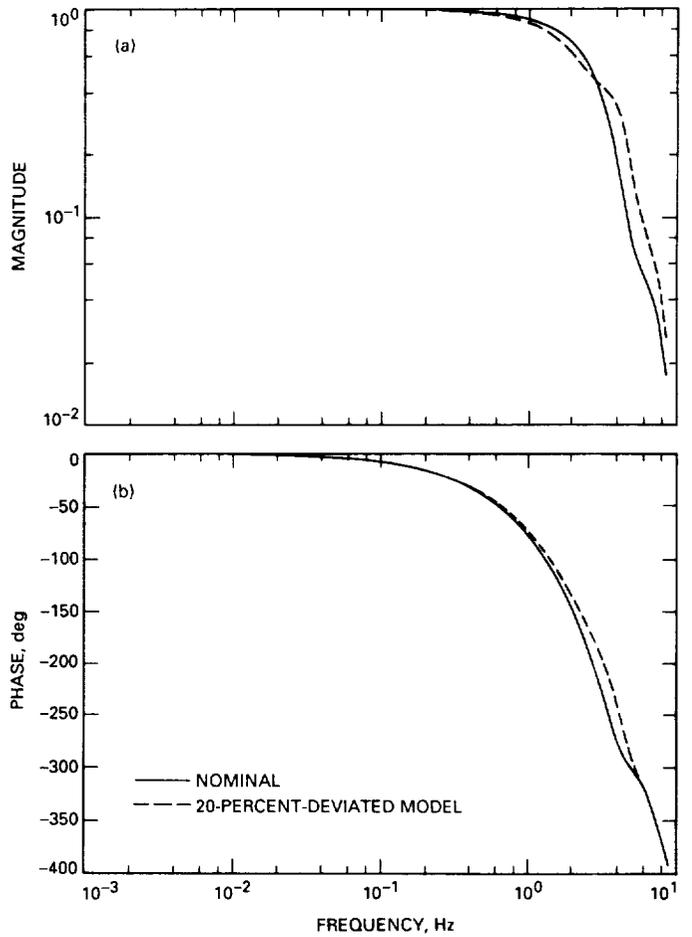


Fig. 22. Closed-loop frequency response: (a) magnitude, and (b) phase.

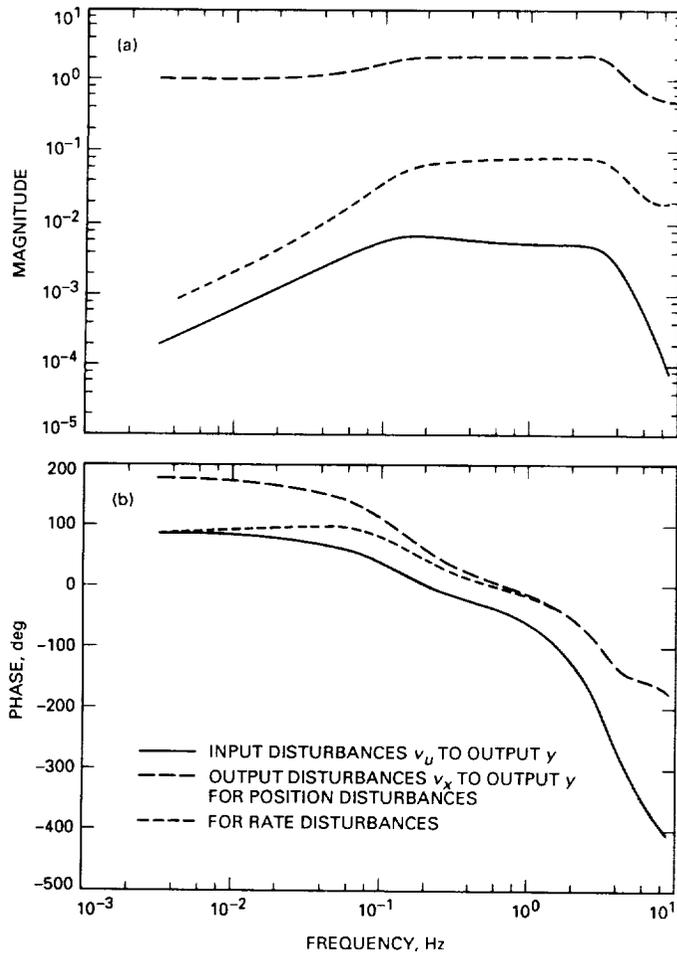


Fig. 23. Transfer functions: (a) magnitude, and (b) phase.

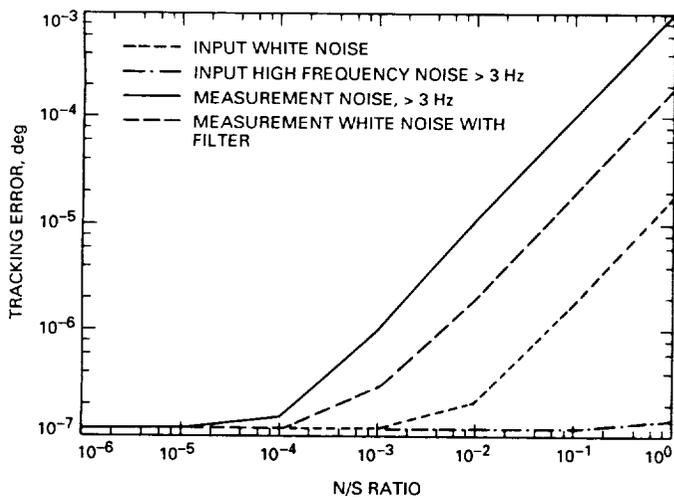


Fig. 24. Tracking errors due to input disturbances.

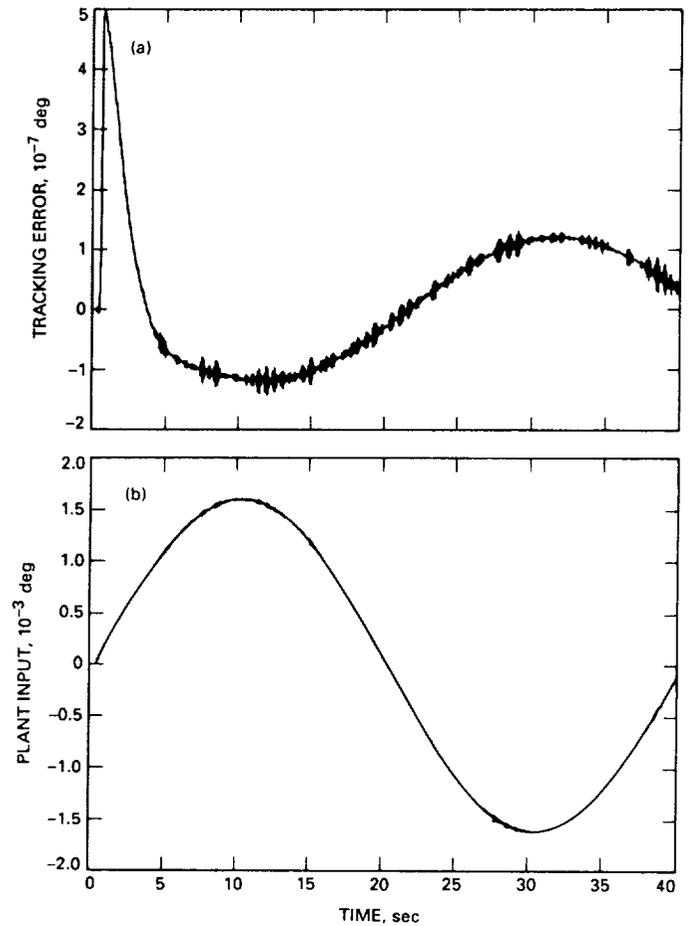


Fig. 25. White noise input disturbances with noise-to-signal ratio = 0.1: (a) tracking error, and (b) antenna input.

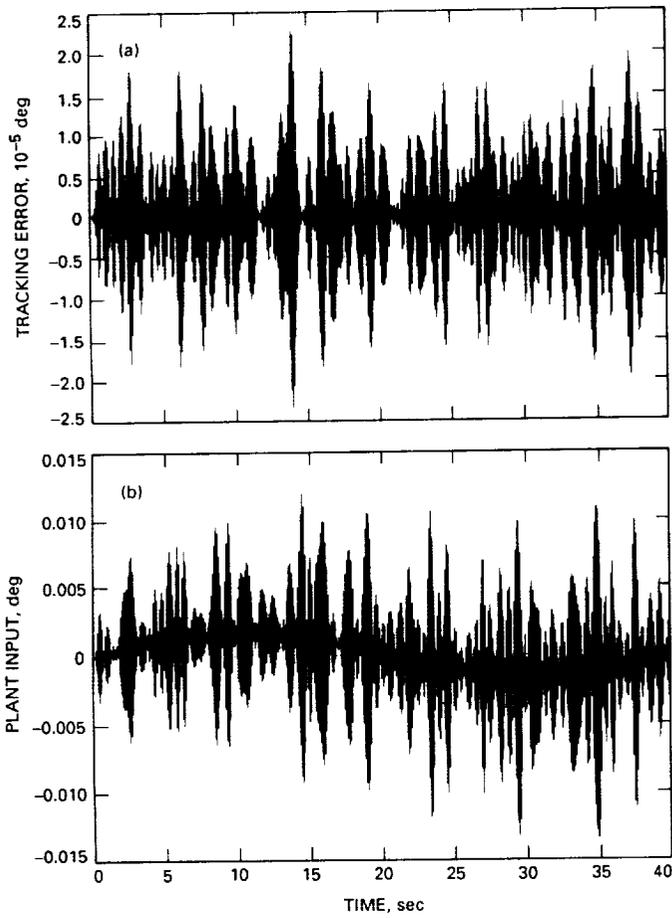


Fig. 26. High-frequency measurement noise with noise-to-signal ratio = 0.01: (a) tracking error, and (b) antenna input.

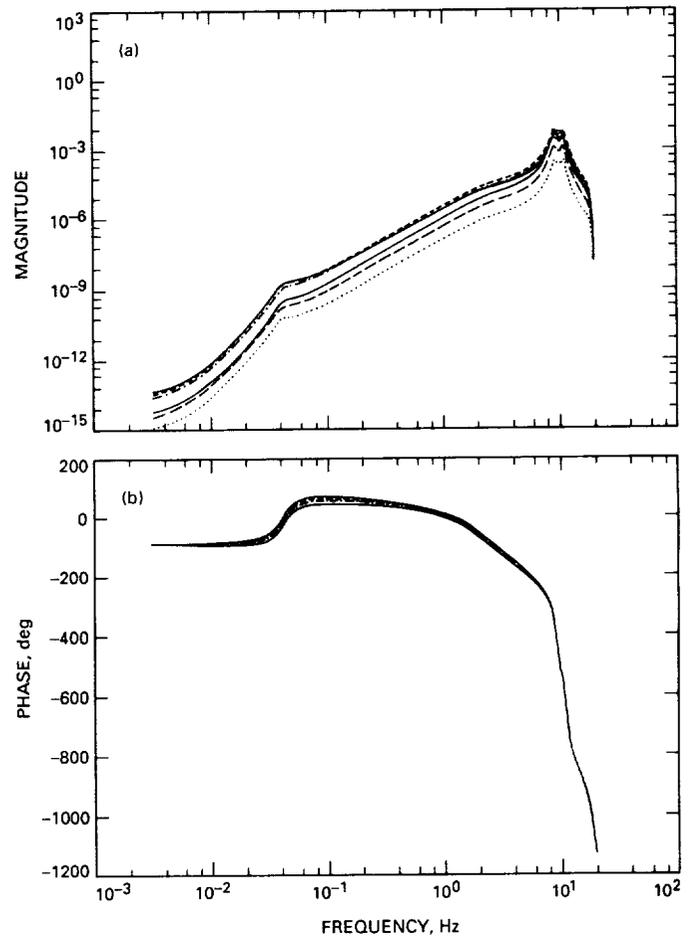


Fig. 27. Transfer functions from output noise  $v_x$  (six states) to output  $y$  for a system with a noise filter: (a) magnitude, and (b) phase.

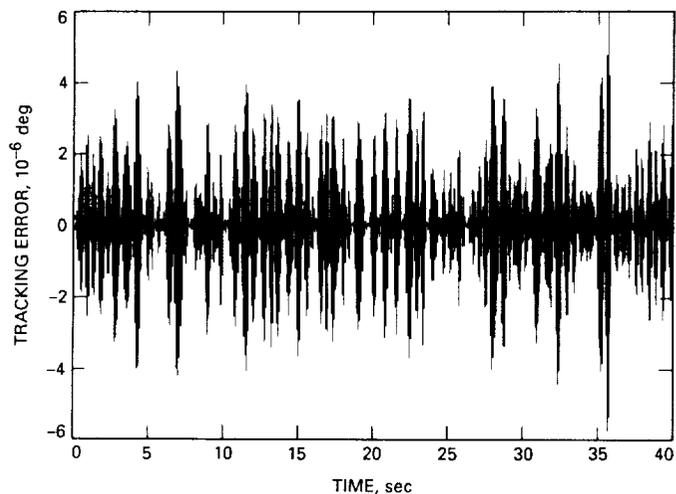


Fig. 28. Tracking error for white-noise output disturbances for a system with a noise filter.

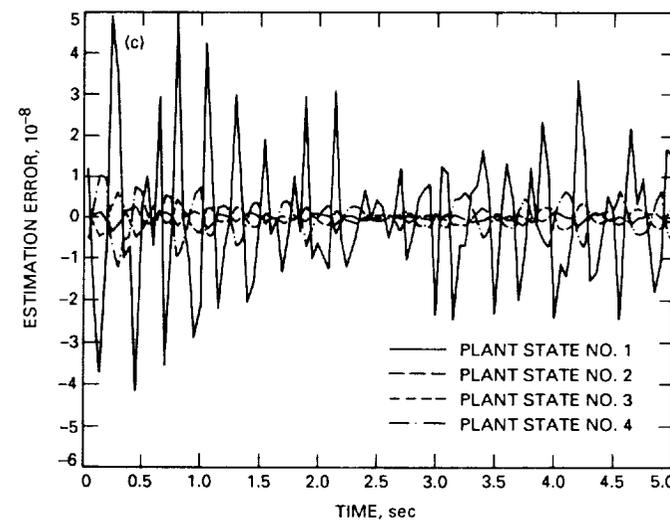
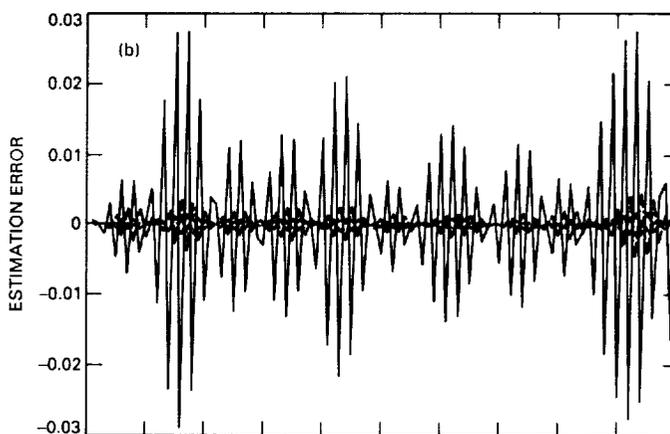
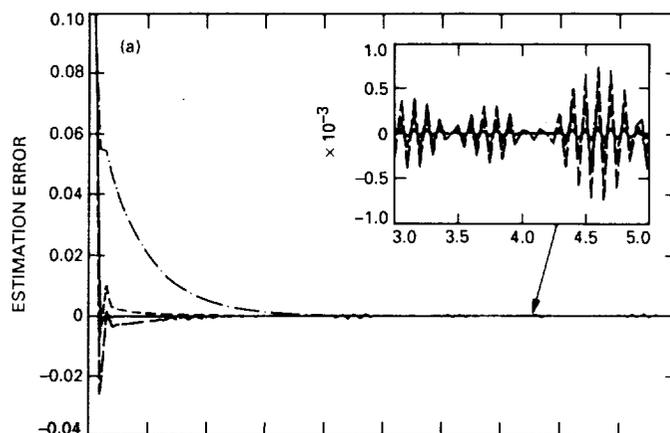


Fig. 30. Estimation error in the presence of measurement noise: (a) LQ estimator; (b) predictive estimator; and (c) predictive estimator with filter.

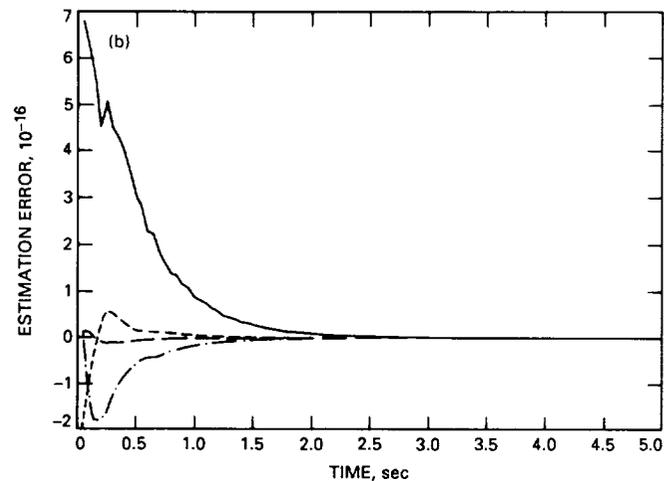
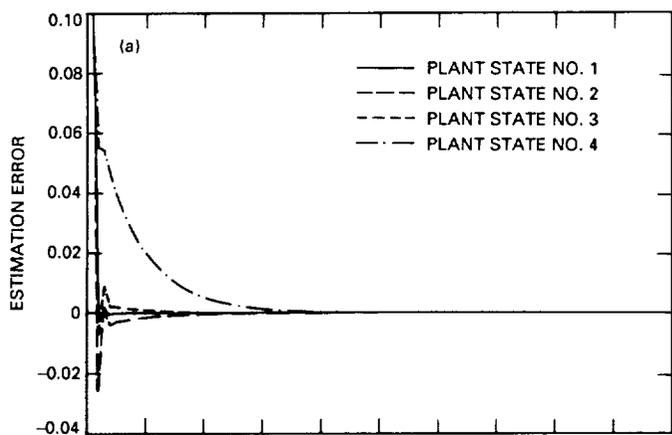


Fig. 29. Estimation error: (a) LQ estimator, and (b) predictive estimator.

55-32

3-22-60

p-18

TDA Progress Report 42-104

February 15, 1991

N91-18313 !

JJ57043h

## Initial Studies of Array Feeds for the 70-Meter Antenna at 32 GHz

P. W. Cramer

Ground Antennas and Facilities Engineering Section

*The results of a study to determine the feasibility of using array feed techniques to improve the performance of the 70-m antenna at 32 GHz are presented. Changing from 8.4 GHz to 32 GHz has the potential of increasing the gain by 11.6 dB, but recent measurements indicate that additional losses of from 3 to 7 dB occur at 32 GHz, depending on the elevation angle. Array feeds have been proposed to recover some of the losses by compensating for surface distortions that contribute to these losses. Results for both surface distortion compensation and pointing error correction are discussed. These initial studies, however, had one significant restriction: The mechanical finite-element model was used to characterize the surface distortions, not the measured distortions from three-angle holography data, which would be more representative of the actual antenna. Further work is required to provide for a more accurate estimate of performance that utilizes holography data and, in particular, one that evaluates the performance in the focal plane region of the antenna.*

### I. Introduction

To achieve a significant performance improvement from the Deep Space Network (DSN), it has been proposed that the operating frequency be increased to 32 GHz. Currently the maximum operating frequency for the 70-m antennas is 8.4 GHz, and these ground antennas were designed to operate efficiently up to this frequency. Assuming that both the spacecraft and ground antennas have the same efficiencies at both frequencies, a performance improvement of 11.6 dB could be expected by changing to 32 GHz. However, measurements performed on the 70-m antenna by Gatti [1] at 32 GHz indicated that losses of 4.6 dB were observed at the rigging angle. The rigging angle is the antenna elevation angle at which the antenna surface

shape has been adjusted to minimize losses. Compared to 1.5 dB at 8.4 GHz [2], this represents a reduction in potential performance by 3.1 dB at the rigging angle.

The 32-GHz measurements also indicated that at an elevation angle of 15 deg the losses were 8.8 dB, an increase of 4.2 dB over the 4.6 dB measured at the rigging angle. At 8.4 GHz the increase in losses over the range of operating elevation angles is under 0.3 dB. The additional losses measured at 32 GHz are a significant part of the potential performance improvement that is to be gained by going to the higher frequency and must be recovered if moving to the higher frequency is to be justified. The work covered by this article analyzes one method that has been proposed to recover a portion of these losses.

Some of these losses are a function of the antenna elevation angle and are due to gravitationally induced reflector surface errors. The surface errors cause distortions in the antenna focal plane field distributions, which are not coupled efficiently into the antenna feed. It has been proposed that an array feed be used to sample the distorted focal plane fields. By proper weighting of the contributions from each array element, the lost performance could then be recovered. This concept would be directed toward recovering a significant portion of the 4.2-dB gravitationally induced surface distortion loss.

A second problem of operating at 32 GHz is the reduction in the antenna pattern beamwidth. The one-sigma pointing accuracy of the 70-m antenna is presently on the order of 0.003 deg. At 32 GHz a pointing error of this magnitude can produce a loss of 1.75 dB. Again, the use of an array feed has been proposed to recover the pointing losses. Blank and Imbriale [3] have analyzed the case for an array feed used with a single distorted parabolic reflector, presenting results covering both distortion compensation and correction for pointing errors. This article extends this work to cover dual-shaped reflector antennas typical of the 70-m antennas used in the DSN.

## II. Analytical Approach

The antenna that is the subject of this study is the 70-m antenna at DSS 14 (Fig. 1). The antenna has optimally shaped reflector surfaces to maximize efficiency. The actual antenna has an offset feed horn and an asymmetric subreflector design to facilitate horn switching. However, to simplify the study, only a non-offset symmetrical geometry was analyzed. Figure 2 shows the layout of the feed horn array, which is located in the focal plane of the antenna. The horns are positioned in a triangular lattice consisting of four rings of elements, the outer boundaries of each ring being hexagonal in shape. This allows the analysis to evaluate different-sized arrays: using one ring with one element, two rings with seven elements, three rings with 19 elements, and four rings with 37 elements. The coordinate system convention used is that the  $z$  axis is along the direction in which the antenna is pointed and the elevation axis is normal to the plane containing the  $y$  and  $z$  axes, with positive  $y$  pointing upward.

To determine the maximum performance that can be achieved with an optimally excited array-fed antenna system, the conjugate weighting method is used [3], as described in the following procedure. The antenna system far field at a common observation point is calculated for

each horn in the array feed. Next, each feed horn excitation is set equal to the conjugate value of the corresponding far-field value at the common observation point. This in effect weights each far-field point by its complex conjugate. Then the weighted far-field values are summed and normalized, as given by

$$G_m = \frac{4\pi \left( \sum_{K=1}^N |E_K|^2 \right)^2}{\eta P_t} \quad (1)$$

and

$$P_t = \frac{1}{\eta} \int_0^{2\pi} \int_0^{\frac{\pi}{2}} |f(\theta, \phi) \sum_{K=1}^N E_K^* e^{jk\hat{R} \cdot \bar{r}_K}|^2 \sin \theta \, d\theta \, d\phi \quad (2)$$

where

$E_K$  = the antenna complex far-field vector associated with the  $K$ th horn

$k$  = wave number

$f(\theta, \phi)$  = field pattern of array horn (assumes no mutual coupling)

$\bar{r}_K$  = array horn position vector in focal plane

$\hat{R}$  = unit vector in direction of observation point

$\theta, \phi$  = spherical coordinates of observation point

$N$  = number of horns in array feed

$\eta$  = free-space impedance

with  $G$  being an estimate of the performance that can be expected from an antenna with distorted surfaces using an array feed to compensate for the errors. To simplify the calculation and to get the best estimate, the direction of the observation point was selected to be in the direction of the pattern peak when only the center element is used.

A finite-element mechanical model [4], identified as model J, is used to describe the antenna geometry and surface shape that results from gravity loading of the antenna at the various elevation angles considered. The errors in the surface shape and antenna geometry were supplied to this task in the form of  $y$ -gravity and  $z$ -gravity load-interpolating coefficients, independent of the elevation angle. To determine the errors at a given elevation angle, the following generic interpolating function is used:

$$P = P_z [\sin(rig) - \sin(elev)] + P_y [\cos(rig) - \cos(elev)] + const \quad (3)$$

where *rig* is the rigging angle and *elev* is the elevation angle of interest,  $P_z$  and  $P_y$  are  $z$  and  $y$  gravity coefficients, and *const* allows the use of a constant term where an error term is not zero at the rigging angle. Each antenna error, such as a subreflector displacement or movement of a reflector surface point, is described by a three-component-vector set of  $P_z$  and  $P_y$ . Using the above expression, the three vector components of a given error  $P$  can be calculated as a function of elevation angle. The mechanical model used assumes that the gravity-induced errors or distortions are symmetrical with respect to the vertical axis ( $y$  axis). Therefore, there are no coefficients for subreflector displacements or antenna beam boresight movements in the  $x$  direction.

The finite-element mechanical model defines the distorted reflector shape in terms of a vector  $\mathbf{V}(u, v, w)$ , which defines a point on the distorted surface relative to a corresponding reference point on a perfect reflector. The antenna pattern analysis program requires that the surface errors be defined by an axial or  $z$ -directed displacement relative to a perfect reflector surface. Figure 3 illustrates the method used to derive the axial term. The vector  $\mathbf{V}(u, v, w)$  is defined relative to the point  $(X_0, Y_0, Z_0)$  on the undistorted main reflector surface. The values of  $X_0$  and  $Y_0$  are used to calculate  $Z_0$ , in this case on the surface of a shaped main reflector. This point, along with vector  $\mathbf{V}(u, v, w)$ , defines a point on the distorted surface  $(X_1, Y_1, Z_1)$ . The value of  $Z_2$  is calculated on the surface of the shaped main reflector at  $(X_1, Y_1)$ . The difference,  $D_z(z)$ , between points  $(X_1, Y_1, Z_1)$  and  $(X_1, Y_1, Z_2)$  defines the axial surface error term.

Since the analysis program needs the distorted surface defined at points other than those in the table of  $(X_1, Y_1, D_z(z))$ , an interpolation function is required. A local interpolating scheme recommended in [5] was selected. The surface of the reflector is subdivided into a number of regions approximately equal to the number of surface panels used on the 70-m antenna. A two-dimensional quadratic function is then best-fitted with up to 16 points from the error table, the points selected being closest to the center of the interpolating region. Fewer than 16 points might be used if the program determines that some of the points are too remote from a given region. The surface error interpolating function is of the form

$$D_z = a_1 + a_2x + a_3y + a_4x^2 + a_5xy + a_6y^2 \quad (4)$$

The procedure used in the analysis is shown in Fig. 4. The gravity load interpolating coefficient table for the reflector, supplied by R. Levy, is the input for the DIST-

RAW program, which computes the actual surface distortions for a given elevation angle. The output of the DIST-RAW program is entered into the DIST-COE program to compute the two-dimensional local interpolating coefficients in Eq. (4) that define the surface errors for the scattering calculation program. A second set of gravity load interpolation coefficients, describing the motion of the subreflector and the antenna beam boresight location, is entered into the RUNGEN program, which calculates the subreflector and boresight location for the specified elevation angle, then calculates the geometry between the reflector surfaces and the array feeds, and finally generates a run stream for calculating the antenna system far-field pattern for each of the array feeds. The horn patterns, the table of data for the undistorted surfaces, and the output of the DIST-COE and RUNGEN programs are entered into the scattering program, and the scattering program is run once for each of the array feed horns. The GTD/Jacobi-Bessel scattering program is used. The output of each run of the scattering program, along with the horn pattern, is entered into the GAIN-EFF program, which evaluates Eqs. (1) and (2).

### III. Optimum Antenna Configuration

To achieve the most improvement and maximize the effectiveness of the array feed in compensating for surface distortions, the subreflector position was adjusted analytically to provide maximum gain at two representative antenna elevation angles, using the calculated pattern for a single standard 22-dB horn. (This capability currently exists, where the DSN antenna subreflectors are moved to compensate for gravity-induced deflections, using a simple elevation-dependent algorithm.) Adjustments were made both vertically and along the antenna axis. At each subreflector position, the gain was calculated in the direction of the predicted boresight angle associated with the selected elevation angle. Using the subreflector adjustments for elevation angles of 15 and 75 deg, a set of gravity coefficients for the Levy interpolation function was derived:

For $z$ -axis movements	$P_y = 0.041850$ in.
	$P_z = -0.941735$ in.
For $y$ -axis movements	$P_y = -2.579083$ in.
	$P_z = -0.545444$ in.

The best boresight location was found to be very close to those predicted by Levy's model J, and therefore his interpolation function coefficients are used for the boresight predictions. The coefficients used are:

Main reflector rotation about $x$ axis	$P_y = -1852.8$ sec $P_z = 0.3$ sec
Main reflector translation in $y$ direction	$P_y = 1332.7$ sec $P_z = 7.0$ sec

The effect of the subreflector translation on the antenna boresight location is not defined in terms of gravity interpolation coefficients. Instead, the following expression supplied by Levy is used:

$$B = 0.0374 Y_t \text{ deg/in. about } x \text{ axis} \quad (5)$$

where  $Y_t$  is the total subreflector motion in the  $y$  direction and  $B$  is the boresight shift contribution due to the subreflector.

In the analysis that follows, it is assumed that the surface errors are strictly due to time-invariant distortions caused by gravitational loads. Any losses due to small-scale surface errors are not accounted for, since these errors are not predicted by the mechanical model. Even if the small-scale surface errors could be predicted, they would not be included because this would require an array with a very large number of unrealizably small elements. In measurements made on the 70-m antenna at 32 GHz and at a rigging angle of 45.5 deg [1], an efficiency of 35 percent was observed for a loss of 4.56 dB. At 8.4 GHz and at an elevation angle of 45 deg, there is a blockage loss of 0.45 dB [2]. If it is assumed that this loss is also typical of the performance at 32 GHz and the calculated directivity efficiency at 45 deg is 0.37 dB, then, subtracting these losses from 4.56 dB, a loss of 3.74 dB remains. Since the efficiency was measured while the antenna was at the rigging angle, where the surface is adjusted to remove any systematic surface errors, it could be assumed that the 3.74-dB loss is due to random small-scale errors in the individual panels. A small-scale error loss of 3.74 dB is equivalent to about a 0.7-mm root mean square (rms) surface error using Ruze's analysis.

Microwave holography imaging at 12 GHz shows all DSN 70-m antennas, in their initial (1988) state of adjustment, as having approximately a 0.7-mm rms error at the rigging angle. Therefore, the assumption that the 3.74-dB loss is a small-scale error loss is reasonable. Since it is not likely that an array feed would be able to compensate for this type of loss, there may be a loss in excess of 3.0 dB which is not recoverable using array techniques. Because the study assumes a model that includes only gravitational loads, this loss does not show up in the following analysis.

If the overall loss is needed, then 4.19 dB (4.56–0.37 dB) needs to be added to the losses or efficiencies presented in this article. In other words, the analysis considered here will not significantly improve the antenna efficiency at the rigging angle, which is about 35 percent as measured in [1]. It should be noted that evidence exists that the 0.7-mm rms small-scale surface errors can be reduced to 0.45 mm by means of a more time-consuming panel adjustment. This procedure might increase the efficiency to approximately 50 percent at the rigging angle.

## A. Effects of Element Size on Performance

Once the best geometry is established, the next step is to determine the effect of the array feed horn element size on the ability of the array to recover lost efficiency. Since the geometry selected places the array elements on a triangular lattice, the element size establishes the element spacing. Four array sizes were evaluated: 1, 7, 19, and 37 elements. An extreme antenna elevation angle of 75 deg was selected for calculation. The results for element diameters ranging from 0.25 in. (0.68 wavelength) to 2.00 in. (5.4 wavelengths) are shown in Fig. 5 for zero-thickness walls. Figure 6 is a similar plot for a smaller range of horn sizes, for horns with 0.05-in.-thick walls. The 0.25-in. diameter is the smallest practical size to be considered, since the cutoff diameter for the  $TE_{11}$  fundamental mode at 32 GHz is 0.216 in. At a diameter of 0.5 in., the  $TM_{11}$  mode can be supported. Therefore, single-mode horns were evaluated for diameters less than 0.5 in., and dual-mode horns were evaluated for diameters of 0.5 in. and larger. Because dual-mode horns equivalent to the 22-dB standard hybrid-mode horns have aperture sizes of about 1.75 in., this size was included as the largest practical size of interest, and the plot was extended to 2.00 in. to see how the curve behaved beyond the largest practical size.

The curve for a single element is what one would expect (Fig. 5). The gain peaks at about 1.75 to 2.00 in., where the best performance would normally be found if no distortions were present, since the antenna optical design was optimized for horns of this size. As the element size becomes smaller, the efficiency drops, as expected, since the antenna reflectors become overilluminated. Looking at the curves for more elements, it can be seen that the additional elements do not compensate for the illumination losses until a diameter of 0.75 in. is reached, and then the performance is still not as good as at 1.75 in. or above. It is not until the single-mode conical horn size of 0.35 in. is reached and 19 or more elements are used that performance equivalent to the larger dual-mode horns is approached. Thus the performance reaches a maximum for horn diameters of approximately 0.35 and 1.75 in.

For a more detailed study, the cases of interest are those with 1.75- and 0.35-in.-diameter horns. To answer the question of whether an intermediate point might be better if pointing or subreflector errors are considered, the 1.25-in. case was also selected. Finally, a case using the standard 22-dB corrugated (hybrid-mode) horns was selected as a reference case. The spacing for the corrugated horns is 2.2 in. to allow space for the corrugations.

The horn wall thickness affects how closely the horns can be positioned. If the horn walls are tapered at the aperture, then they can be spaced as if there were no wall thickness. This is the case in Fig. 5. To illustrate the effect of wall thickness, a set of calculations for horns with 0.05-in. walls at the aperture was made, and the results are plotted in Fig. 6. The efficiency for a series of horns with 0.05-in.-thick walls falls off more rapidly than for zero thickness simply because the horn size refers to the maximum horn diameter and the wall thickness then detracts from the horn's effective aperture. The effect is more pronounced for the smaller sizes, since the wall thickness accounts for a larger percentage of the horn size. For all arrays with elements smaller than 1.75 in., zero wall thickness gives significantly better results. For 1.0-in.-diameter elements, for example, the differences are on the order of 1.0 dB. This is one illustration of how critical the feed design is to achieving the maximum recovery of the energy in the antenna focal region.

## B. Effects of Antenna Elevation Angle

Figures 7 through 10 show the performance of the array feed as a function of the antenna elevation angle. (Note the change in the range of elevation angles in Fig. 8.) It can be seen that the performance peaks at 45 deg and then drops off as the elevation angle either increases or decreases. The best performance is obtained at 45 deg because the antenna surface shape is adjusted at this angle to compensate for any gravitational distortions. This is referred to as the rigging angle. The rigging angle could be any angle, but for the structural model used in this study the angle was set at 45 deg. As the antenna elevation angle diverges from the rigging angle, further gravitational errors cause the surface to deviate from the optimum shape and the antenna efficiency begins to degrade. With the exception of the case shown in Fig. 8, the elevation angles for this study range from 7.5 to 75 deg, which covers the operational angles imposed on the antenna system.

The effects of element spacing can be seen by comparing the curves for the standard hybrid-mode horn (2.2-in. diameter) in Fig. 7 and the 1.75-in. dual-mode horn in Fig. 8. For the single-element case the performance is the same for

both horn types to within a few hundredths of a decibel over the range of elevation angles. This is to be expected, since the patterns for a single horn for these two cases are very similar. The efficiency ranges from  $-3.5$  through  $-0.4$  to  $-2.3$  dB over these elevation angles, showing that there is significant room for improvement. When seven elements are used, the dual-mode horns with their denser packing have better performance by 0.5 dB at 7.5 deg than the hybrid-mode horns. With 37 elements, the difference at 7.5 deg is 0.6 dB, and at 75 deg it is 0.13 dB.

Considering the 1.75-in. dual-mode configuration only, the following observations can be made. At 45 deg, the loss is 0.42 dB for a single horn, which is the directivity loss for an undistorted reflector; with a properly designed feed such as the one used, it represents the best performance achievable. Using additional horns at 45 deg does not change the performance. At an elevation angle of 7.5 deg, a single element has a loss of 3.48 dB, seven elements have a loss of 2.04 dB, and 37 elements have a loss of 1.72 dB. Thus there is an improvement of 1.44 dB when going from one to seven elements and 1.77 dB when going to 37 elements. At an elevation angle of 75 deg, a single element has a loss of 2.35 dB, seven elements have a loss of 1.73 dB, and 37 elements have a loss of 1.49 dB. In this case there is an improvement of 0.62 dB when going from one to seven elements and 0.85 dB when going to 37 elements. The array feed in effect has halved the losses due to surface distortion when using 37 elements at the extremes in elevation angles, with most of the improvement achieved by adding one ring of elements for a total of seven elements.

The performance curves shown in Fig. 9 for the 1.25-in. dual-mode horns are considerably lower than those for the previous cases at all elevation angles and for any number of horns. The loss at 45 deg with a single element is 1.33 dB because the antenna is overilluminated by the smaller horn. It is interesting to note, however, that with additional elements at this elevation angle no significant performance improvement is achieved. This shows that, for shaped-reflector designs and with small reductions in horn size, the outer horns do not effectively capture the small amount of energy no longer collected by the center element. For the 0.35-in. single-mode horn (Fig. 10), it takes 19 elements to get to within 0.0 to 0.2 dB of a single 1.75-in. horn over the range of elevation angles calculated. It takes 37 elements to get within 0.1 dB of a single 1.75-in. horn at a 45-deg elevation angle. Thus, for cases where the antenna is properly pointed and the subreflector is in the optimum position for the elevation angle, the smaller elements provide no real advantage and the 1.75-in. dual-mode horn is the better choice.

### C. Performance Versus Pointing Angle

In addition to compensating for distortions with a properly pointed antenna, an array feed can also be used to correct for antenna pointing errors. These pointing errors can be due to rapid changes in the surface distortion, a lack of knowledge of what the pointing errors are for a given elevation angle, or a limitation in the pointing accuracy of the antenna control system. A pointing accuracy on the order of 0.003 deg is considered good for the 70-m antenna and causes a negligible loss at 8.4 GHz, the current maximum operating frequency. At 32 GHz, however, a pointing error of this magnitude can give rise to a 1.75-dB loss. Pointing errors were simulated in this study by calculating the performance improvements along directions at various angles relative to the best antenna pointing angle (boresight). With one exception, the calculations were made at an elevation angle of 75 deg so as to include the effects of surface distortions. Cases were calculated for 0.35-in. single-mode horns, for 1.25- and 1.75-in. dual-mode horns, and for a 2.2-in. hybrid-mode horn. The one exception is the 2.2-in. hybrid-mode horn, for which an elevation angle of 45 deg was used.

The performance of the 2.2-in. hybrid-mode horn is shown in Figs. 11 and 12. For an elevation angle of 45 deg (Fig. 11), it can be seen that the performance as a function of boresight offset angle is virtually the same for any number of elements. This indicates that only one element is contributing to the antenna performance. Figure 13 illustrates the antenna beam patterns for four feed elements (one located on the antenna axis and three located at various radial distances from the axis). For elements other than the central one of an array feed to compensate for losses due to a pointing error, they must contribute signal power in the direction of the pointing error. If a pointing error of 0.003 deg is assumed, from the figure it can be seen that only the first beam has any energy in that direction. The next beam ( $Y = 1.8621$ ) is considerably more than 16 dB down from the contribution of the center element at 0.003 deg. (Note that the curves do not extend low enough to provide a more accurate value.) At this level the second element (or for that matter any of the other additional elements) cannot significantly contribute to improving the performance for a pointing error on the order of 0.003 deg. This effect can be seen for a 75-deg elevation angle in Fig. 12, where the separation of the four curves remains essentially constant as a function of boresight offset angle over typical pointing errors.

As will be seen later, significant pointing error improvements will not be seen until the feed size, and therefore the feed spacing, is reduced by three or more times and a larger

number of elements is used. This problem is aggravated by the shaped reflector design of the 70-m antenna. Shaped designs provide nearly uniform illumination, which in turn provides the narrowest beamwidth for a given antenna size. Conventional antenna designs do not provide uniform illumination and therefore have much wider beamwidths. These wider beamwidths for conventional-design antennas allow higher crossovers between adjacent beams and therefore a higher potential to compensate for pointing errors than for shaped designs.

The following discussion applies to elevation angles of 75 deg. The hybrid-mode horn case (Fig. 12), over the range of  $-0.004$  to  $0.002$  deg (for a pointing accuracy of 0.003 deg) with seven elements, gives an improvement over and above the distortion compensation of only 0.2 dB at  $-0.004$  deg and 0.0 dB at  $0.002$  deg. Additional elements do not give any additional pointing-error compensation. The 1.75-in. dual-mode horn case (Fig. 14) has a performance very similar to the hybrid-mode horn case over the range of  $-0.004$  to  $0.002$  deg, except for a 0.2-dB improvement at  $0.002$  deg for seven elements. Thus the larger horn sizes show little potential for pointing correction. The 1.25-in. dual-mode horn case (Fig. 15) shows improved pointing capability with additional elements over the same angles. Unfortunately, because of the lower performance at the best pointing angle, all the element curves for the 1.25-in. case are below the corresponding ones for the 1.75-in. case over a range of pointing errors of 0.006 deg. Therefore the 1.25-in. case is not as good a choice as the larger array sizes unless it is a requirement to support a range of pointing errors larger than 0.008 deg.

The arrays with smaller elements have much better pointing-error correction capability. The correction capability for the 0.35-in.-diameter single-mode horn is shown in Fig. 16. The performance with one or seven elements is not as good as can be achieved with 1.75-in. elements, since with such a small number of elements the antenna is overilluminated. The performance of the 0.35-in. case with 37 elements and no pointing correction falls between that for one element and that for seven elements for the 1.75-in. case. Over a range of pointing errors of 0.006 deg, the 0.35-in. case with 37 elements has from 1.0 to 1.2 dB better performance than the 1.75-in. case with seven elements and 0.8 and 0.9 dB better performance than the 1.75-in. case with 19 elements. The 0.35-in. case with 19 elements has 0.2 to 0.3 dB better performance than the 1.75-in. case with seven elements and no improvement over the 1.75-in. case with 19 elements. For error ranges greater than 0.006 deg, which could result from wind gusts, 0.35-in. arrays with 19 or more elements have

an advantage. This advantage, however, is at the expense of the better performance that would be achieved with 19 1.75-in. horns if the pointing errors were normally limited to a range of errors smaller than 0.006 deg.

For pointing error compensation and with 37-element designs, arrays with 0.35-in.-diameter horns are best. For 19-element designs, arrays with 0.35-in.-diameter horns are better when pointing errors are greater than  $\pm 0.003$  deg. For seven-element designs, arrays with 1.75-in.-diameter horns are better, even though the 1.75-in. design has no pointing correction capability.

#### D. Performance Versus Subreflector Lateral Position

The ability of the array feed to improve performance was analyzed with the subreflector located at the position that gives the best performance as a function of elevation angle. In actual practice, the subreflector will normally be set at this location, and it is useful to know how accurately the subreflector needs to be positioned to maintain optimal performance. In addition, knowledge of the effect of array design on subreflector positional accuracy would be useful in the selection of an optimum array design. First consider the case using the hybrid-mode horn. Figure 17 shows the antenna at 45 deg, and therefore the results do not include any distortion effects. The number of elements has no effect on the performance until the positional errors exceed 0.1 in., well within the expected positional accuracy. In order to hold the losses within approximately 0.2 dB, the subreflector needs to be set within 0.03 in. Figure 18 shows the performance at an elevation angle of 75 deg. Although the performance is lower because of the surface distortions, the sensitivity to subreflector position for a single element remains about the same. Over the region of interest, additional elements do not improve performance. Figures 19 through 21 show the performance of the smaller feed elements. The 1.75-in. dual-mode design behaves essentially the same as the hybrid-mode design over the small range of displacements expected. The relative performance between elements of different sizes is similar to that seen for beam-pointing performance. The subreflector is less sensitive to position for the 1.25-in. dual-mode case than for the 1.75-in. case when the use of more than one element is considered. The 0.35-in. single-mode case has a low sensitivity to position with 19 elements, the minimum number of elements that is practical with elements of this size. With 37 elements, the 0.35-in. design is virtually insensitive to subreflector position over  $\pm 0.1$  in. However, if the subreflector can be positioned within  $\pm 0.04$  in., including wind gusts, then the seven-element 1.75-in. horn case still represents the best overall performance.

## IV. Summary and Conclusions

The previous discussion covers the calculated performance of the 70-m antenna when used beyond its design frequency, where surface distortions detract from its performance. An array of circular feed horns arranged in a triangular lattice was used to recover some of the lost performance. A study was conducted to determine how effective this method is in recovering the lost performance so that the potential improvement from increasing the operating frequency can be judged.

Although the cost tradeoff for this concept is beyond the scope of this study, some useful observations can be made. It was indicated that the measured gain loss at 32 GHz at the rigging angle was 4.56 dB with a hybrid-mode feed horn. The calculated radio frequency (rf) losses, which do not include random small-scale surface distortions, quadripod blockage, or dissipation losses, for a single hybrid-mode horn with the antenna at the rigging angle are 0.37 dB. This indicates that the small-scale surface distortion, quadripod blockage, and dissipation losses are 4.19 dB. Adding this loss to the calculated rf loss of 3.48 dB at an elevation angle of 7.5 deg gives an overall loss of 7.67 dB. This is to be compared with a 1.8-dB loss at 8.4 GHz, composed of a 1.5-dB efficiency loss [2] and an antenna distortion loss of 0.3 dB at 7.5 deg. This gives a net loss of 5.87 dB at 32 GHz relative to 8.4 GHz at 7.5 deg. In order to evaluate the significance of this relative loss, this loss along with any relative losses associated with the spacecraft system would have to be subtracted from the potential gain increase of 11.6 dB due to frequency scaling. Thus, subtracting the 5.87-dB relative loss from 11.6 dB leaves a 5.73-dB net improvement, less any spacecraft-related losses.

As mentioned earlier, the small-scale rms error of the 70-m antennas is about 0.7 mm. A large part of this error is panel setting error. Recent data indicate that the panels could be readjusted, using holography techniques, to provide an rms error of about 0.45 mm. If this improvement can be achieved, 2.2 dB of the small-scale surface error losses could be recovered and as much as a 7.93-dB net improvement could be obtained. A long-term goal is to achieve a 0.25-mm rms surface accuracy by the year 2000, and this could provide 1.0 dB of additional improvement at 32 GHz.

As the study shows, some of the lost performance can be recovered with an array feed. It was further shown that arrays with 1.75-in.-diameter horns gave the best improvement for distortion errors. For seven elements at 7.5 deg, the array feed would recover 1.44 dB of the estimated 3.48-dB gravity distortion loss. Going to 19 or

37 elements would recover only an additional few tenths of a decibel and is not practical to consider at this time.

To provide pointing-error compensation, small array elements are required, with the 0.35-in.-diameter horn elements performing best. Although the peak gain of the 1.75-in. array elements in the boresight direction is higher than for the 0.35-in. elements, the scanning capabilities of a 37-element array of 0.35-in. elements allow it to exceed the static performance of an array of 1.75-in. elements. For pointing errors in excess of 0.003 deg, an array of 19 0.35-in. elements will give better performance. Using fewer than 19 elements for the 0.35-in. element array is not practical, since they would cover a smaller area than a single 1.75-in. horn. Since a single 1.75-in. horn is close to optimum for an undistorted antenna, a 0.35-in. array with fewer than 19 elements would never be able to effectively illuminate an undistorted antenna, much less a distorted one.

The question naturally arises as to why more of the energy lost to large-scale surface distortions cannot be recovered. Why is the recovery limited to about half of the lost power? There is no simple answer at this time. The analysis was done in the transmit mode. Patterns and excitations were assigned to the array feed, and the ef-

fect on the overall gain of the antenna was calculated in the presence of surface errors. This method gives no insight into what the focal fields look like and how they are affected by the antenna distortions. In addition, the transmit method gives no idea of how effectively the array feed system samples these focal plane fields. What is required is to calculate the performance in the receive mode. This would directly provide the focal plane fields. In turn, by correlating the aperture fields of the array horns with the focal plane fields, the performance of the antenna can be predicted. More importantly, greater visibility of what is happening to the fields in the focal plane region would provide a method for determining the best array geometry and horn type for surface-error compensation.

A program capable of efficiently computing the receive-mode patterns of a 70-m antenna at 32 GHz did not exist at the time this study began. Also the field correlation technique required needed to be developed. Further work needs to be done to implement and use the receive mode of analysis, and then other classes of feed horn designs need to be studied to determine how much more of the lost power can be recovered. If more of the lost power is recoverable, then it must be determined what the potential improvements are and whether array feeds represent the best way to implement a 32-GHz capability.

## Acknowledgment

The author expresses appreciation to Dr. Roy Levy for supplying the mechanical model, along with the tables of gravity coefficients from which the distortion data required for this study were derived.

## References

- [1] M. S. Gatti, M. J. Klein, and T. B. H. Kuiper, "32-GHz Performance of the DSS-14 70-Meter Antenna; 1989 Configuration," *TDA Progress Report 42-99*, vol. July–September 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 206–219, November 15, 1989.
- [2] D. A. Bathker and S. D. Slobin, "DSN 70-Meter Antenna Microwave Optics Design and Performance Improvements; Part I: Design Optimization," *TDA Progress Report 42-97*, vol. January–March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 306–313, May 15, 1989.
- [3] S. J. Blank and W. A. Imbriale, "Array Feed Synthesis for Correction of Reflector Distortion and Vernier Beamsteering," *IEEE Trans. Antennas Propagation*, vol. AP-36, no. 10, pp. 1351–1358, October 1988.
- [4] R. Levy and R. Melosh, "Computer Design of Antenna Reflectors," *Journal of the Structural Division, ASCE*, vol. 99, no. ST11, Proc. Paper 10178, pp. 2269–2285, November 1973.
- [5] V. Galindo-Israel, W. A. Imbriale, Y. Rahmat-Samii, and T. Veruttipong, "Interpolation Methods for Shaped Reflector Analysis," *IEEE Trans. Antennas Propagation*, vol. AP-36, no. 3, pp. 441–444, March 1988.

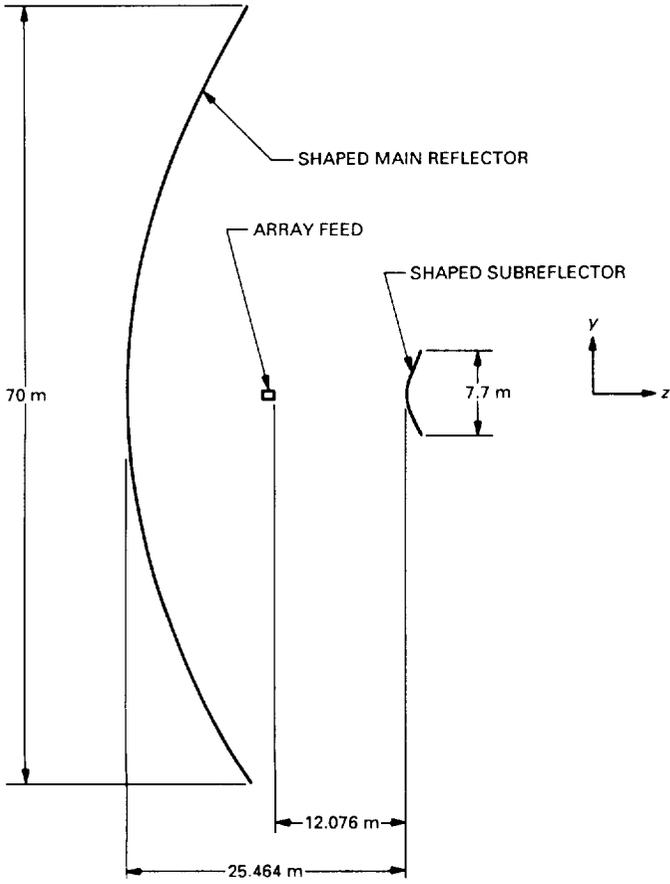


Fig. 1. Geometry of 70-m antenna.

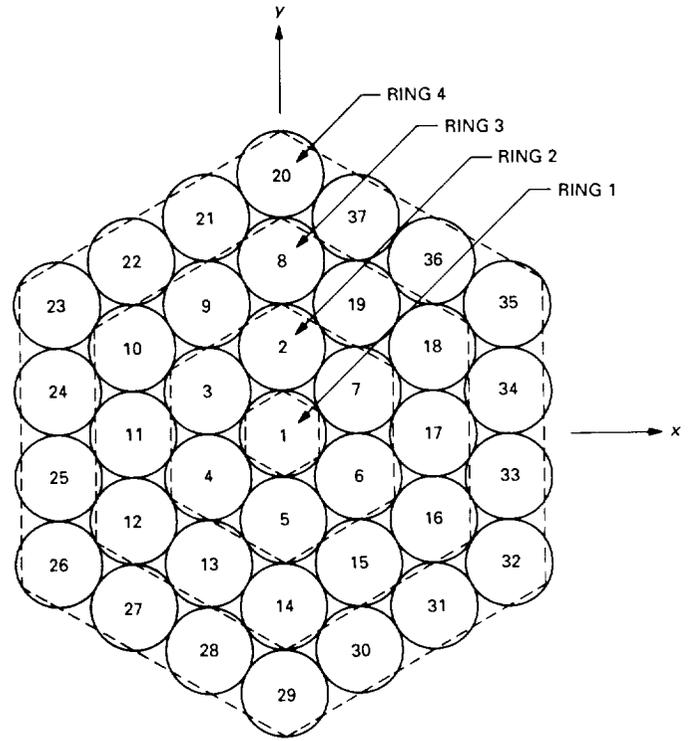


Fig. 2. Feed array configuration.

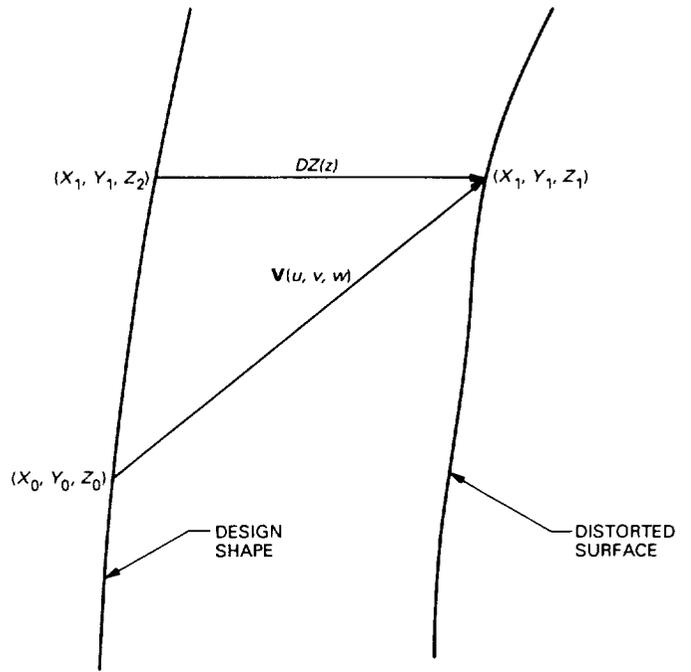


Fig. 3. Definition of surface distortion.

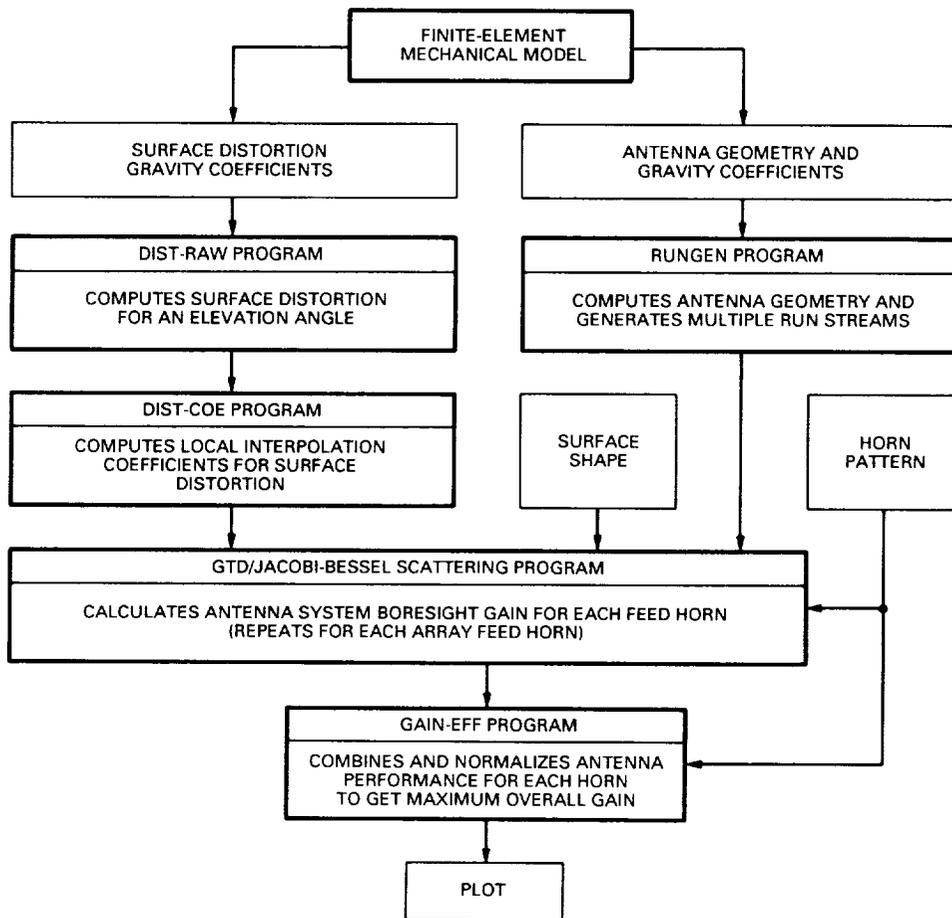


Fig. 4. Calculation flow diagram for analysis.

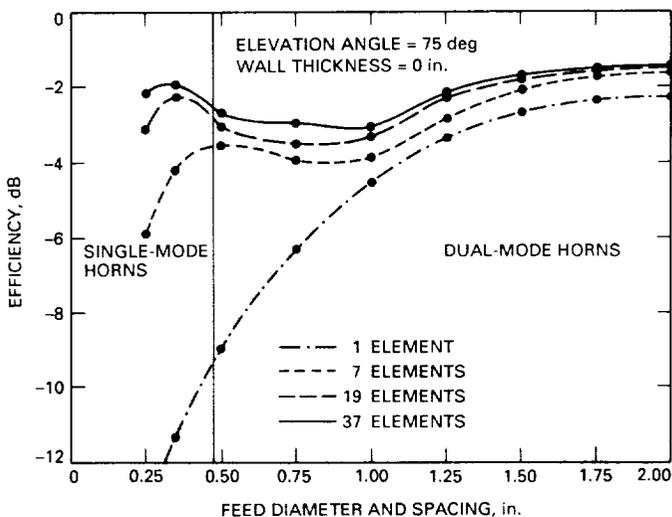


Fig. 5. Effect of feed element size on efficiency of 70-m antenna for arrays of elements with zero-thickness walls.

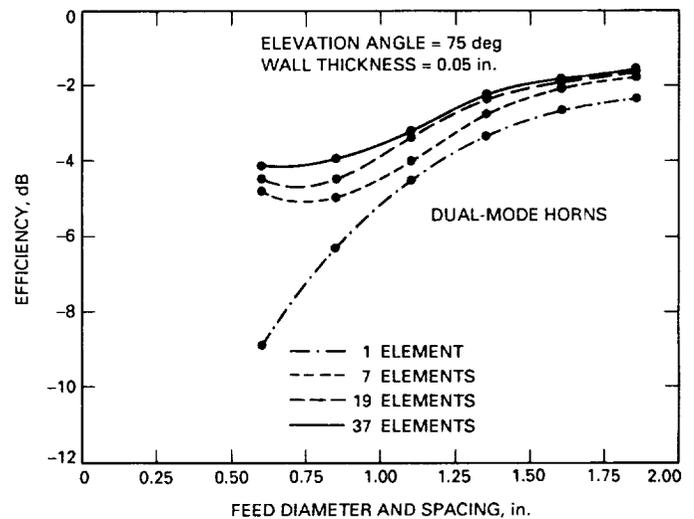


Fig. 6. Effect of feed element size on efficiency of 70-m antenna for arrays of elements with 0.05-in.-thick walls.

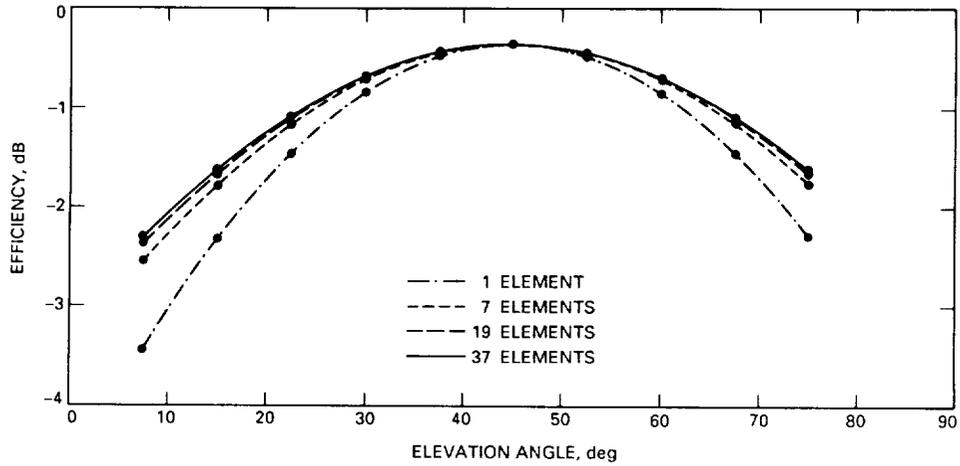


Fig. 7. Performance of 70-m antenna versus antenna elevation angle for arrays of 2.2-in.-diameter hybrid-mode feed horns.

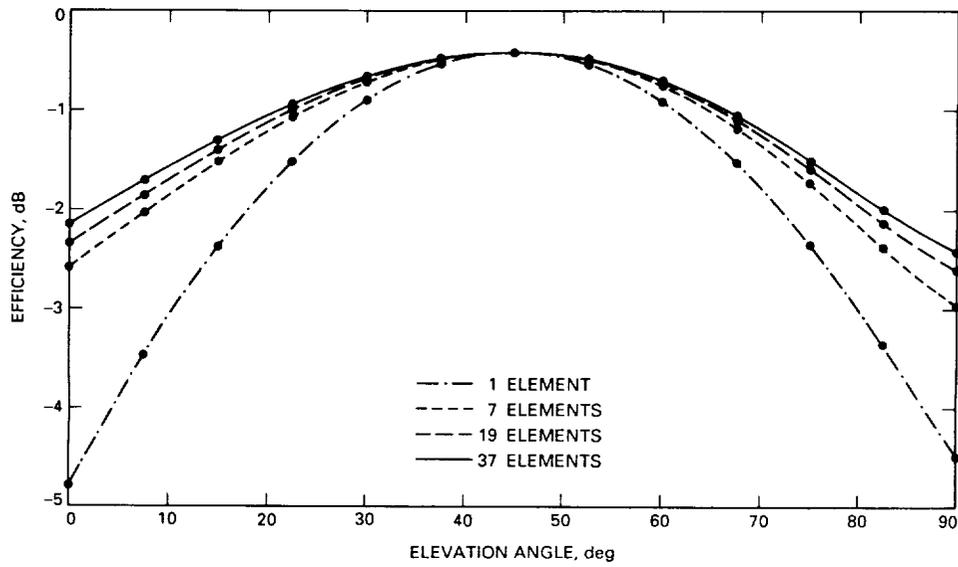


Fig. 8. Performance of 70-m antenna versus antenna elevation angle for arrays of 1.75-in.-diameter dual-mode feed horns.

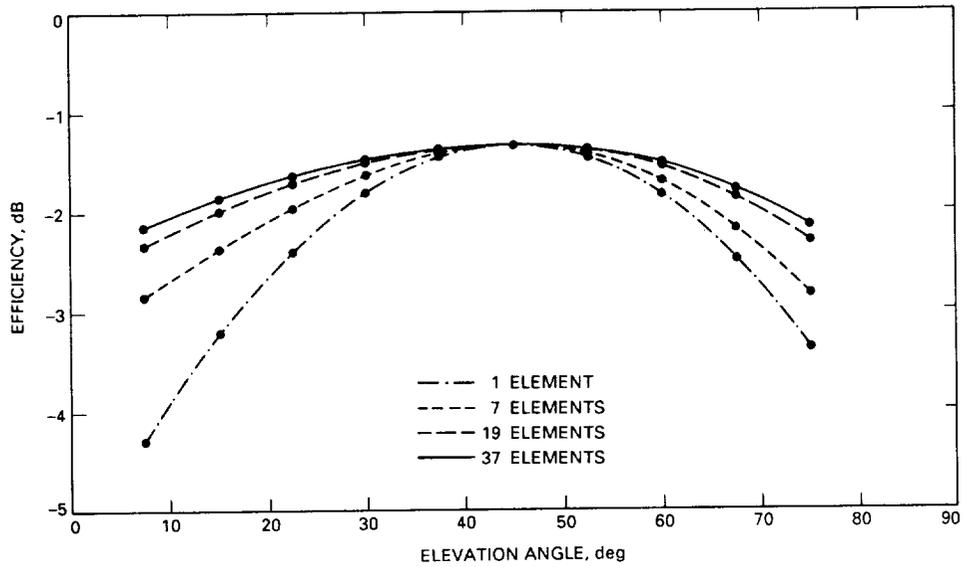


Fig. 9. Performance of 70-m antenna versus antenna elevation angle for arrays of 1.25-in.-diameter dual-mode feed horns.

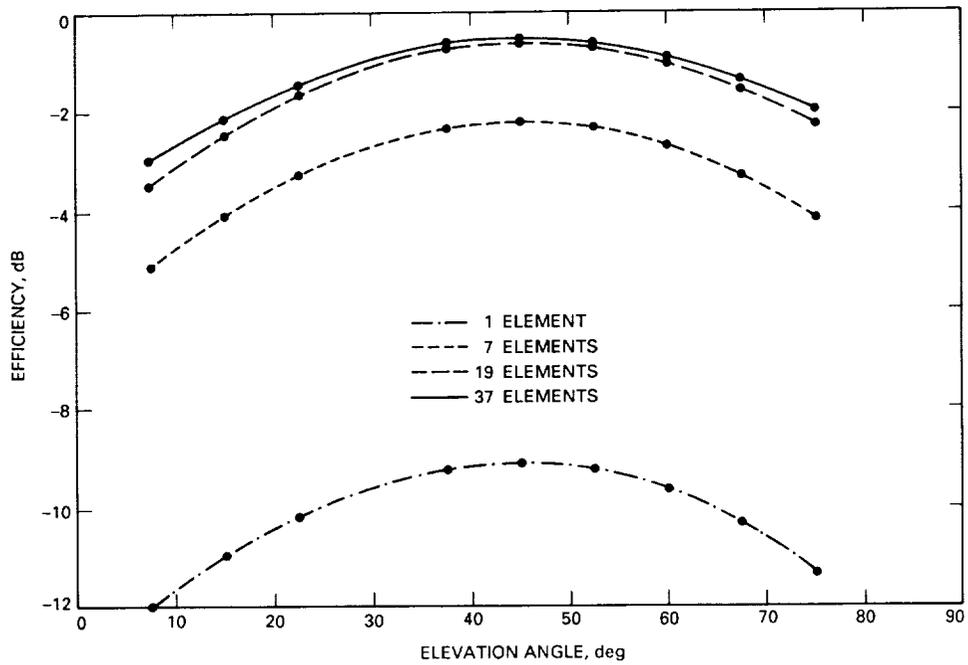


Fig. 10. Performance of 70-m antenna versus antenna elevation angle for arrays of 0.35-in.-diameter single-mode feed horns.

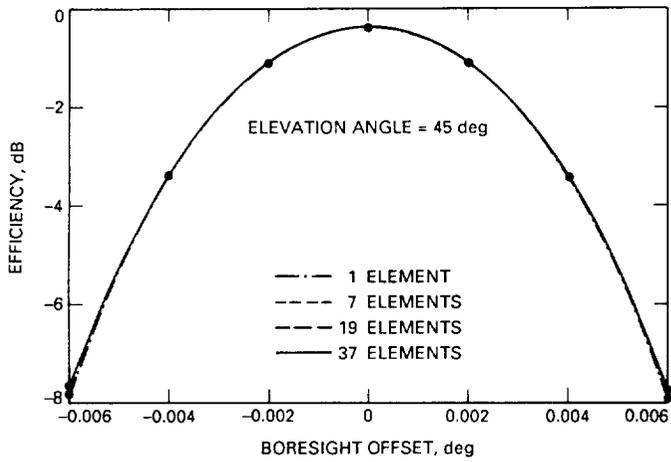


Fig. 11. Effect of antenna pointing error on performance of 70-m antenna for arrays of 2.2-in.-diameter hybrid-mode feed horns at 45 deg.

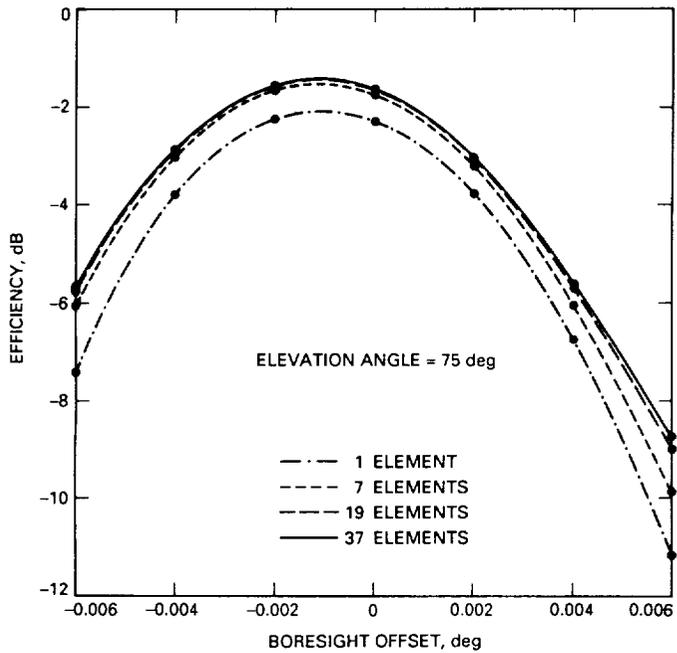


Fig. 12. Effect of antenna pointing error on performance of 70-m antenna for arrays of 2.2-in.-diameter hybrid-mode feed horns at 75 deg.

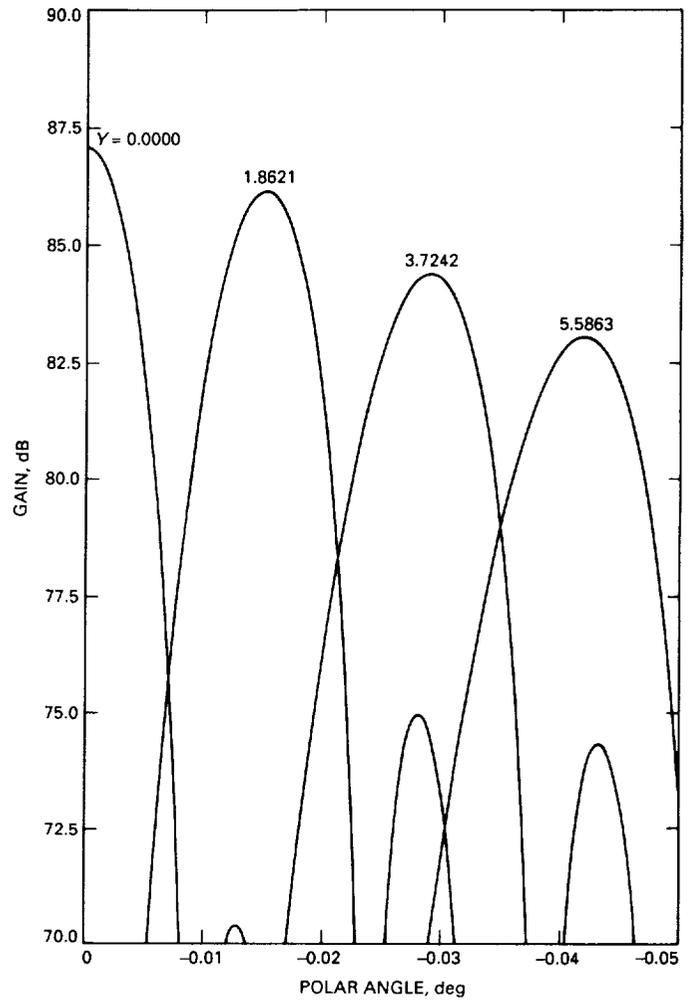


Fig. 13. Beam patterns of 70-m antenna for hybrid-mode feed horns at four different lateral offset positions.

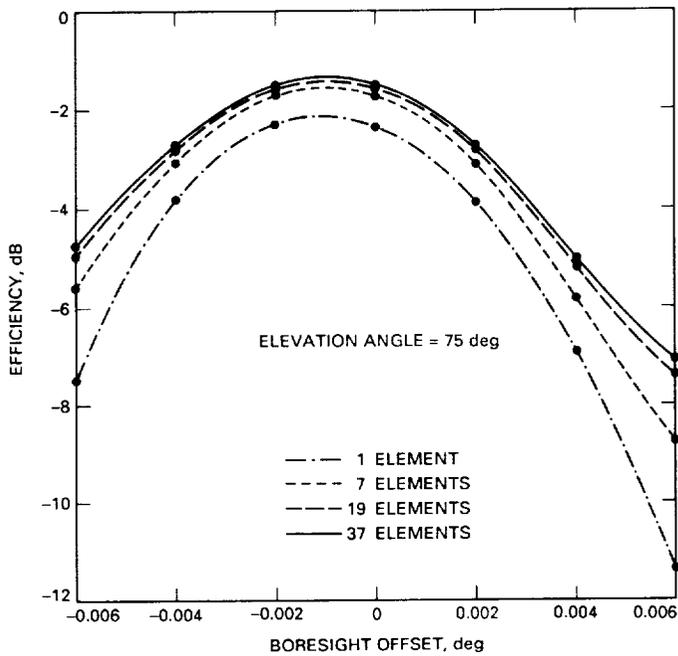


Fig. 14. Effect of antenna pointing error on performance of 70-m antenna for arrays of 1.75-in.-diameter dual-mode feed horns.

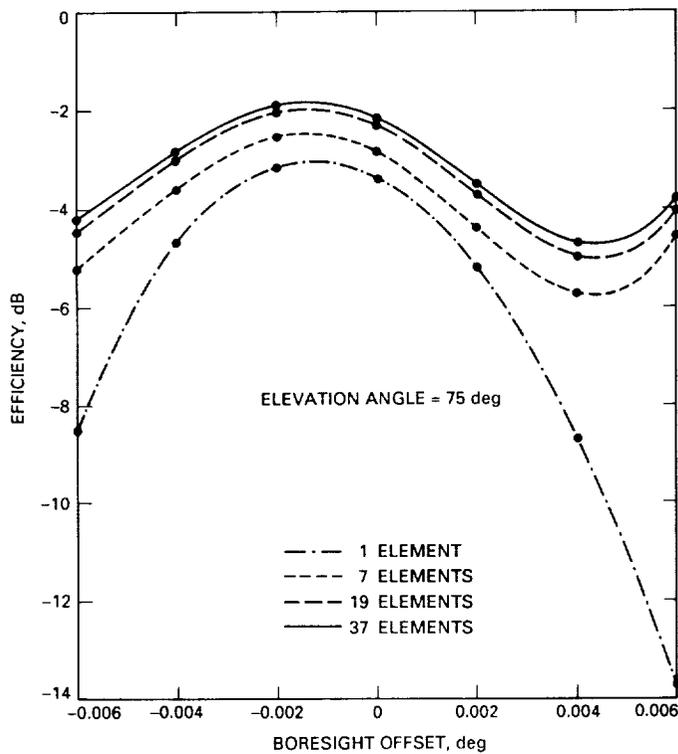


Fig. 15. Effect of antenna pointing error on performance of 70-m antenna for arrays of 1.25-in.-diameter dual-mode feed horns.

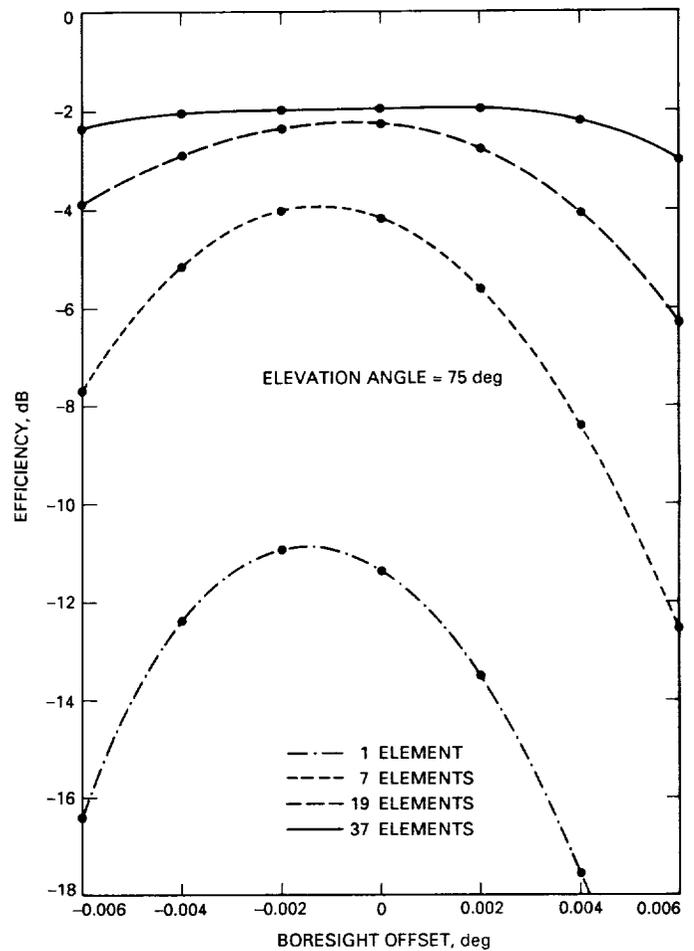


Fig. 16. Effect of antenna pointing error on performance of 70-m antenna for arrays of 0.35-in.-diameter single-mode feed horns.

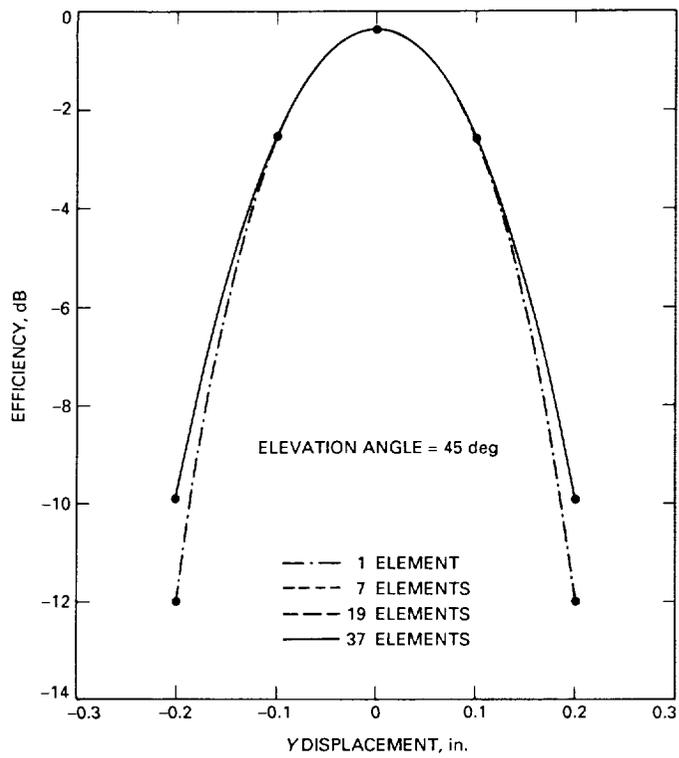


Fig. 17. Effect of subreflector displacement on performance of 70-m antenna for arrays of 2.2-in.-diameter hybrid-mode feed horns at 45 deg.

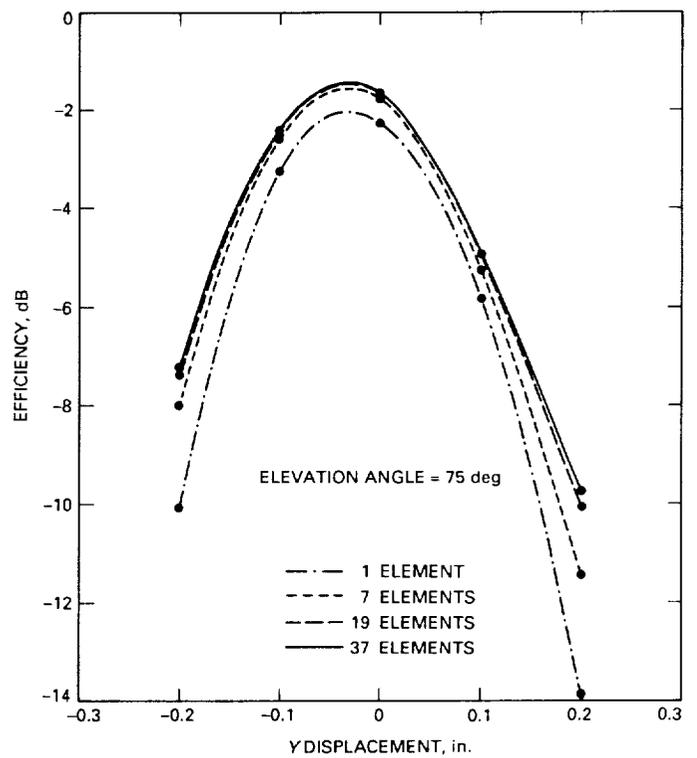


Fig. 18. Effect of subreflector displacement on performance of 70-m antenna for arrays of 2.2-in.-diameter hybrid-mode feed horns at 75 deg.

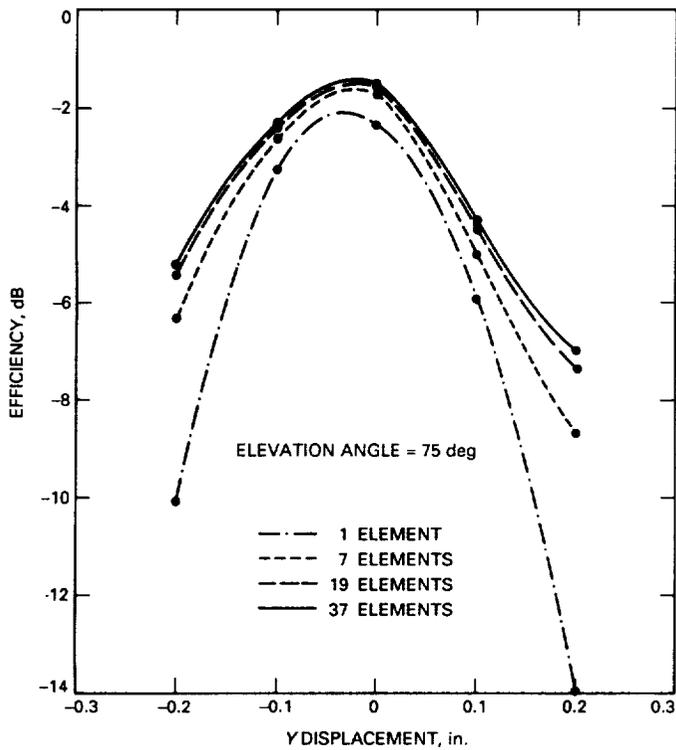


Fig. 19. Effect of subreflector displacement on performance of 70-m antenna for arrays of 1.75-in.-diameter dual-mode horns.

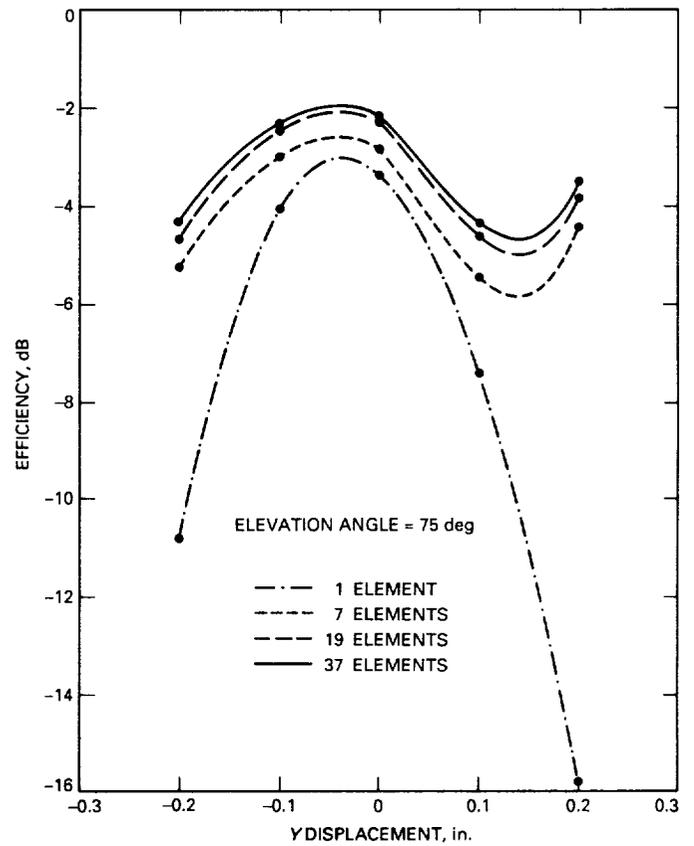
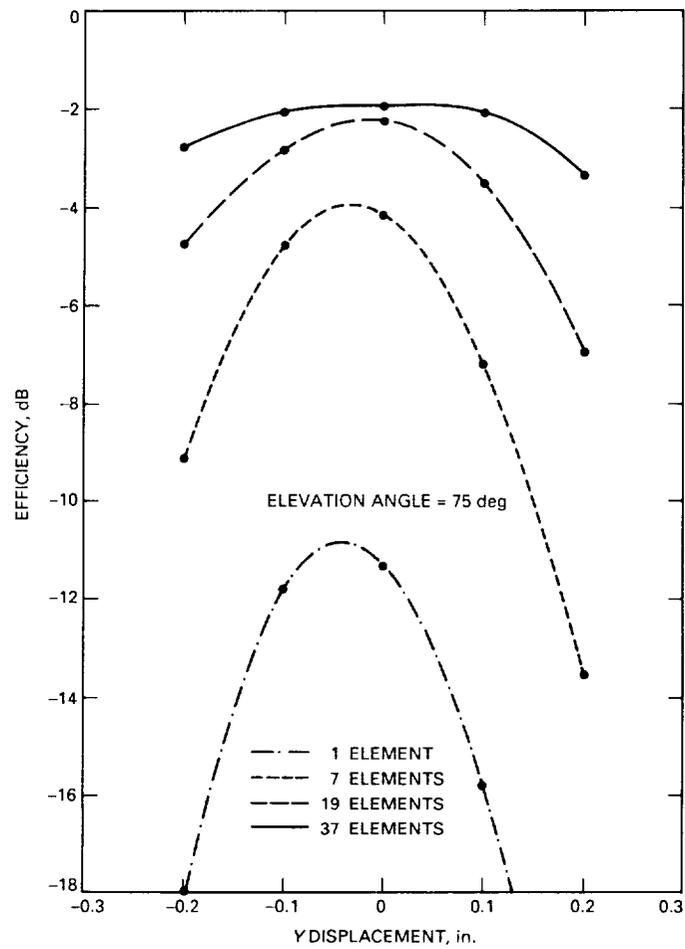


Fig. 20. Effect of subreflector displacement on performance of 70-m antenna for arrays of 1.25-in.-diameter dual-mode horns.



**Fig. 21. Effect of subreflector displacement on performance of 70-m antenna for arrays of 0.35-in.-diameter single-mode horns.**

56-32

332  
11-10

**N91-18314**

JJ574450

## **Steady-State Phase Error for a Phase-Locked Loop Subjected to Periodic Doppler Inputs**

C.-C. Chen and M. Z. Win

Communications Systems Research Section

*The performance of a carrier phase-locked loop (PLL) driven by a periodic Doppler input is investigated. By expanding the Doppler input into a Fourier series and applying the linearized PLL approximations, it is easy to show that, for periodic frequency disturbances, the resulting steady-state phase error is also periodic. Compared to the method of expanding frequency excursion into a power series, the Fourier expansion method can be used to predict the maximum phase-error excursion for a periodic Doppler input. For systems with a large Doppler-rate fluctuation, such as an optical transponder aboard an Earth-orbiting spacecraft, the method can be applied to test whether a lower order tracking loop can provide satisfactory tracking and thereby save the effort of a higher order loop design.*

### **I. Introduction**

Coherent carrier phase recovery using a phase-locked loop (PLL) has become an integral part of digital communication systems [1-3]. By performing coherent demodulation using the recovered signal carrier, the receiver can achieve 3 dB of performance gain over systems using non-coherent demodulation techniques. The ability to recover and track the incoming carrier phase can also lead to a significant performance gain in related applications such as coherent ranging and spacecraft navigation [4].

The performance of a phase-locked receiver depends critically on the ability to accurately recover the carrier

phase. Synchronization errors between the incoming signal and the local reference can quickly lead to a degraded signal-to-noise ratio (SNR) and a large power penalty. The design of the loop, therefore, must ensure proper phase tracking under the operating conditions. In general, the performance of the PLL is influenced by the additive circuit noise, the oscillator frequency noise, and the frequency characteristics of the signal it is designed to track. The effect of channel noises on the performance of the PLL has been studied extensively [1-4]. It is shown that, in general, the residual phase tracking error due to the additive noise increases with PLL bandwidth, whereas the tracking error due to the oscillator frequency noise decreases with increasing loop bandwidth. With a given SNR, therefore,

there is an optimal choice of receiver bandwidth and PLL design that minimizes the root-mean-square (rms) phase tracking error.

In addition to the channel and oscillator noises, the performance of the PLL is also affected by the frequency characteristics of the source. A simple first-order loop can provide adequate tracking only when the free-running frequency of the voltage-controlled oscillator (VCO) is equal to the frequency of the incoming signal. A second-order loop is needed to track a constant frequency offset, whereas a third-order loop is needed to track a signal with linearly varying frequency. In general, higher order loops are needed to compensate for higher order frequency disturbances. However, higher order loops present extra design complications since control loops higher than second order are not unconditionally stable. Furthermore, for systems with large dynamic frequency fluctuations, the frequency perturbations with orders higher than the control loop cannot be completely compensated by the loop. As a result, some residual phase tracking error always remains.

Since the performance of the PLL-based receiver depends on the ability to accurately track the carrier phase, it is important to estimate the amount of residual tracking error due to the higher order frequency perturbations. The analysis of PLL performance given a time-dependent Doppler input is generally accomplished by expanding the Doppler signal into a power series and then analyzing the effects of different power terms separately. The power-series expansion provides a simple and intuitive method of expanding the Doppler frequency term. However, it is difficult to justify dropping the higher order terms since the steady-state phase error due to higher order terms does not converge.

For a very special class of system, the Doppler signal is periodic. In this case, the phase-locked loop equation can be examined by performing a Fourier decomposition rather than a power-series expansion. In this article, the procedure of analyzing the residual phase tracking error using a Fourier expansion of the frequency perturbation is outlined.

## II. Carrier Phase Tracking Loop

The essentials of a carrier PLL include a phase detector, a loop filter, and a VCO. Shown in Fig. 1 is a typical implementation of a radio-frequency (RF) PLL. The phase detector detects the phase difference between the incoming signal and the output of the VCO. For carrier phase tracking applications, an RF mixer is generally used as

the phase detector. The mixer output is filtered by the loop filter with transfer function  $F(s)$ . The sum frequency term at the mixer output is filtered by the loop filter such that the loop effectively responds only to the difference frequency term. When the frequency of the signal is equal to that of the VCO, the difference frequency term is simply proportional to  $\sin \phi(t)$ , where  $\phi(t)$  is the phase difference between the signal and VCO output.

The filtered phase-difference signal is subsequently injected into the receiver VCO. The output frequency of the VCO is linearly dependent on the input voltage signal. When the loop eventually reaches a locked condition, the phase-error signal,  $\phi(t)$ , will be such that it is governed by the following loop equation:

$$\frac{d\phi(t)}{dt} + AKf(t) \otimes [n(t) + \sin \phi(t)] = \beta(t) + f_N(t) \quad (1)$$

where  $n(t)$  is the additive noise,  $f_N(t)$  is the oscillator frequency noise, and  $\beta(t)$  is the frequency error between the signal and the local oscillator. The loop mechanization is represented by the signal amplitude,  $A$ ; the VCO gain constant,  $K$ ; and the impulse response of the loop filter,  $f(t)$ . For systems operating with an ample signal-to-noise ratio, the effect of additive noise is usually very small. At the same time, a relatively high SNR allows oscillator frequency noise to be tracked out. For this analysis, therefore, the focus is on the frequency detuning term.

If the phase error is small (the loop is in lock), the sine function can be approximated by its argument, and the phase error of the PLL can be adequately described using the following linearized form:

$$\frac{d\phi(t)}{dt} + AKf(t) \otimes \phi(t) = \beta(t) \quad (2)$$

The integral-differential equation in Eq. (2) can be simplified into a linear differential equation of the form

$$\begin{aligned} \frac{d^n \phi(t)}{dt^n} + \dots + a_1 \frac{d\phi(t)}{dt} + a_0 \phi(t) = \\ b_{n-1} \frac{d^{n-1} \beta(t)}{dt^{n-1}} + \dots + b_1 \frac{d\beta(t)}{dt} + b_0 \beta(t) \end{aligned} \quad (3)$$

where the coefficients  $\{a_j\}$  and  $\{b_j\}$  are related to the loop transfer function  $H(s)$  by

$$\frac{b_{n-1}s^{n-1} + \dots + b_1s + b_0}{s^n + \dots + a_1s + a_0} = \frac{1}{s + AKF(s)}$$

$$= \frac{1}{s}[1 - H(s)] \quad (4)$$

Equivalently, it can also be simplified into a set of first-order differential equations [5]. Given the frequency detuning process,  $\beta(t)$ , Eq. (3) can be solved for the steady-state phase error. For a time-varying frequency detuning process, one method of simplifying the analysis is to expand  $\beta(t)$  into a Taylor series and then retain only terms sufficient for the analysis. Since the system is linear, the solution obeys the superposition principle and is equal to the sum of solutions of individual expansion terms.

In general, the solution to Eq. (3) includes the transient response and the steady-state response terms. The transient response, which depends on the initial conditions, dies out after a time period that is on the order of the inverse loop bandwidth. If the phase error converges to a constant, the steady-state solution can be easily solved by Laplace transforming Eq. (3) and using the final value theorem, i.e.,

$$\lim_{t \rightarrow \infty} \phi(t) = \lim_{s \rightarrow 0} s \cdot \Phi(s) = \lim_{s \rightarrow 0} \frac{sB(s)}{s + AKF(s)} \quad (5)$$

where  $F(s)$  and  $B(s)$  are the Laplace transforms of  $f(t)$  and  $\beta(t)$ , respectively. Equation (5) is applicable only when the steady-state solution exists as a constant value. For higher order perturbations, the final value in Eq. (4) does not converge, and the Laplace transform cannot be used to solve for the steady-state response. Fortunately, it is known from the linear differential equation theory that the general solution to Eq. (3) for a driving force of the form  $\beta(t) = a_n t^n$  is a polynomial of order  $n$ . The steady-state phase error can therefore be solved by substituting the polynomial of order  $n$  into the right-hand side of Eq. (3) and then matching the coefficients.

Although the steady-state phase error can be solved by assuming a polynomial general solution, the resulting polynomial is diverging at  $t \rightarrow \infty$ . Since most physical systems do not have unbounded frequency variation, the higher order perturbation eventually dies down. The analysis of the PLL performance using higher order perturbation is therefore limited to the time period within which the perturbation is present. The loop design is said to be adequate if the effects of the higher order perturbation are small. Because of the complexity of designing higher order

tracking loops, the analysis of the time-varying Doppler term is generally limited to third order or less. Justifications for dropping the higher order terms, however, can be very difficult since the solution is not bounded.

### III. Fourier Expansion of the Time-Dependent Doppler Signal

Since the frequency fluctuation at the input is generally bounded, polynomial approximation to the Doppler signal will eventually become greater than the input. If the time period of interest is longer than the time for the polynomial approximation to deviate from the signal, a better (higher order) approximation is needed to analyze the PLL behavior. For some class of missions such as Earth-orbiting satellites, however, the periodic orbit will result in a periodic Doppler input that should intuitively result in a periodic phase variation. For such a system, the steady-state solution can be more easily derived by expanding the Doppler signal into a Fourier series. The resulting linearized PLL equation can be written as

$$\frac{d\phi(t)}{dt} + AKf(t) \otimes \phi(t) = \beta(t) = \sum_{k=-\infty}^{\infty} c_k e^{ik\omega_0 t} \quad (6)$$

where  $\omega_0$  is the fundamental frequency (reciprocal of the period) of the perturbation. From linear differential equation theory, it is known that a linear differential equation responding to a sinusoidal driving term with frequency  $\omega_0$  will exhibit a general solution with an identical frequency. Again, the particular solution (transient response) is expected to die down with a time constant that corresponds to the eigenvalues of the characteristic equation. Furthermore, by using the superposition principle, solutions to different harmonics can be solved individually. Consequently, if the Doppler stimulus can be expanded into a Fourier series, the solution can be found using the superposition principle.

Two examples can now illustrate the power of this technique.

**Example 1: First-Order Loop.** It is well known from linearized PLL theory that the first-order loop can be used to track a constant frequency detuning (Doppler) with a constant phase offset. Higher order perturbation can result in a loss of lock. If such a loop is used to track a periodic frequency variation of frequency  $\omega_0$ , conventional analysis cannot adequately predict the resulting loop performance. However, by performing the spectral expansion of the Doppler signal

$$\beta(t) = c_0 \sin \omega_0 t \quad (7)$$

the general solution to the first-order loop can be written as

$$\phi(t) = \frac{a_0 c_0}{a_0^2 - \omega_0^2} \sin \omega_0 t + \frac{\omega_0 c_0}{a_0^2 - \omega_0^2} \cos \omega_0 t \quad (8)$$

where  $a_0 = AK = 4B_L$  is related to the bandwidth of the loop. It is seen from Eq. (7) that a first-order loop can be used to track a periodic Doppler input, provided that the loop bandwidth, frequency variation, and Doppler period satisfy the condition for linearizing the loop equation, i.e.,  $\phi(t) \ll 1$  for all  $t$ . Furthermore, the resulting steady-state phase error is periodic with the same period as the driving term, but falls slightly out of phase from the driving input.

**Example 2: Perfect Second-Order Loop.** It is known from the linearized loop theory that a perfect second-order loop can be used to track out a constant Doppler rate with a steady-state phase error. The transfer function for the loop filter is

$$F(s) = \frac{1 + \tau_2 s}{\tau_1 s} \quad (9)$$

By substituting Eq. (9) into the PLL equation and converting the resulting equation back to the time domain,

$$\frac{d^2 \phi(t)}{dt^2} + a_1 \frac{d\phi(t)}{dt} + a_0 \phi(t) = \frac{d\beta(t)}{dt} \quad (10)$$

where  $a_1 = AK\tau_2/\tau_1$ , and  $a_0 = AK/\tau_1$ . The general solution to this equation can be reached by substituting the solution of the form

$$\phi(t) = P \sin \omega_0 t + Q \cos \omega_0 t \quad (11)$$

into Eq. (10) and equating the coefficients. It is found that the general (steady-state) solution due to the periodic Doppler input is given by

$$\begin{aligned} \phi(t) = & \frac{c_0 \omega_0^2 a_1}{(a_0 - \omega_0^2)^2 + a_1^2 \omega_0^2} \sin \omega_0 t \\ & + \frac{c_0 \omega_0 (a_0 - \omega_0^2)}{(a_0 - \omega_0^2)^2 + a_1^2 \omega_0^2} \cos \omega_0 t \end{aligned} \quad (12)$$

Again, the constants must satisfy the constraints that  $\phi(t) \ll 1$ . Note that the loop bandwidth can be related to the constants  $a_0$  and  $a_1$  by

$$B_L = \frac{a_0}{4a_1} + \frac{a_1}{4} \quad (13)$$

The difference between the Fourier solution and the power-series solution can be seen in Fig. 2 where the steady-state phase response of the linear loop to a sinusoidal frequency excursion has been plotted. The power-series solution was calculated by expanding the sinusoid into a power series and retaining the first two terms (Doppler rate and second derivative of Doppler). It is seen from the figure that the power-series solution is a close approximation to the actual solution during the initial 1/4 period. However, as soon as the approximation to the sinusoid breaks down, the power-series estimate diverges, whereas the actual solution remains bounded.

## IV. Discussion

Expanding the Doppler signal into a Fourier series offers a different perspective in predicting the PLL performance. Unlike the power-series expansion method that, although intuitive, cannot adequately predict the performance under higher order perturbations, the periodic expansion of the Doppler results naturally in a periodic phase solution. As a result, the theory can predict a bounded solution even when the driving force (Doppler) has a higher order component. For applications where a good phase synchronization is essential, such as coherent data communications, the Fourier solution can provide an adequate estimate of the maximum phase-error excursion.

An example for the problem occurs in the design of the optical phase tracking loop between a low Earth-orbiting satellite and a ground station. At the operating wavelength of 1  $\mu\text{m}$ , the relative Doppler rate of the two terminals can change from +300 MHz/sec to -300 MHz/sec within 30 seconds. Given a PLL of 20 kHz bandwidth, it is difficult to predict whether the loop can remain adequately in lock during the period. By approximating the Doppler signal near the portion of the orbit with maximum frequency change as a sinusoid, however, a simple estimate of the PLL performance can be obtained. For a frequency excursion with 3-GHz amplitude and 60-second period, it can be shown that the loop can adequately track the Doppler with less than 0.2 radian of residual error.

Although the Fourier expansion provides a bounded solution for a periodic varying Doppler signal, there are some

practical limitations in applying the Fourier analysis technique. First, the solution presented above ignores the transient solution. This is true only when the period of the dynamic signal is long compared to the loop-response time. Furthermore, the results are derived only for a linearized equation. For a nonlinear PLL equation, a periodic driv-

ing force can excite higher harmonic terms. Finally, for a predictable periodic driving force, it is a usual practice to apply a periodic estimator correction term at the VCO input to compensate for the periodic driving force. In this case, the detuning is small and the loop will essentially respond only to the noise inputs.

## References

- [1] W. C. Lindsey and M. K. Simon, *Telecommunications Systems Engineering*, Englewood Cliffs, New Jersey: Prentice-Hall, 1973.
- [2] A. J. Viterbi, *Principles of Coherent Communications*, New York: McGraw-Hill, 1966.
- [3] J. J. Spilker, *Digital Communication by Satellite*, Englewood Cliffs, New Jersey: Prentice-Hall, 1977.
- [4] W. C. Lindsey, *Synchronization Systems in Communication and Control*, Englewood Cliffs, New Jersey: Prentice-Hall, 1972.
- [5] C. T. Chen, *Introduction to Linear System Theory*, New York: Holt, Rinehart and Winston, 1970.

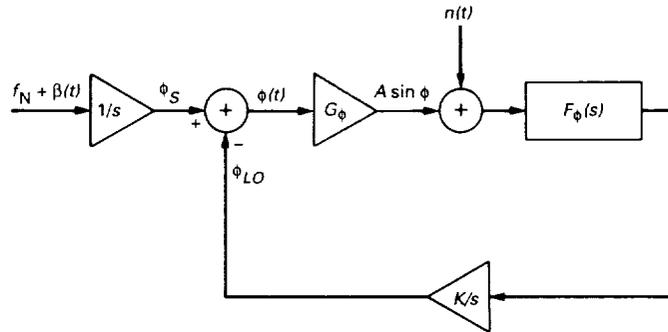


Fig. 1. Equivalent-noise block diagram of a carrier phase-locked loop.

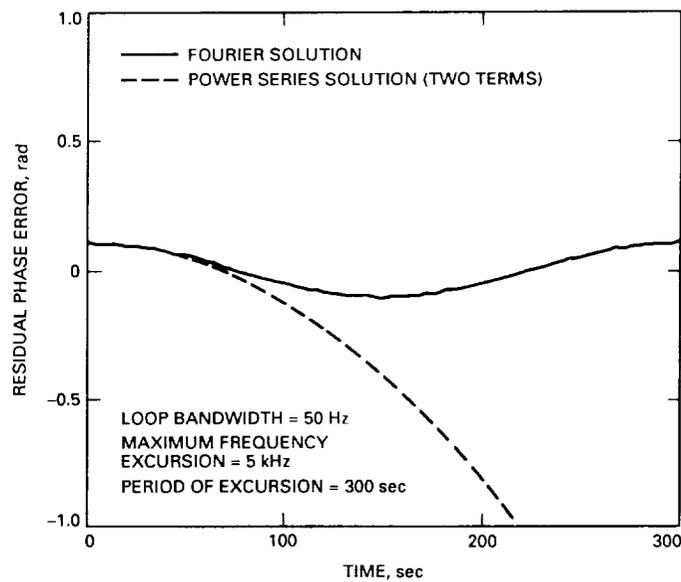


Fig. 2. Residual phase tracking error for a second-order phase-locked loop subjected to a periodic Doppler input. The PLL was analyzed by using the Fourier decomposition method and a power-series solution.

57-65

p-14

**N91-18315**

JJG/MS

## **Maximum-Entropy Probability Distributions Under $L_p$ -Norm Constraints**

S. Dolinar

Communications Systems Research Section

*This article tabulates continuous probability density functions and discrete probability mass functions which maximize the differential entropy or absolute entropy, respectively, among all probability distributions with a given  $L_p$ -norm (i.e., a given  $p$ th absolute moment when  $p$  is a finite integer) and unconstrained or constrained value set. Expressions for the maximum entropy are evaluated as functions of the  $L_p$ -norm. The most interesting results are obtained and plotted for unconstrained (real-valued) continuous random variables and for integer-valued discrete random variables.*

*The maximum entropy expressions are obtained in closed form for unconstrained continuous random variables, and in this case there is a simple straight-line relationship between the maximum differential entropy and the logarithm of the  $L_p$ -norm. Corresponding expressions for arbitrary discrete and constrained continuous random variables are given parametrically; closed-form expressions are available only for special cases. However, simpler alternative bounds on the maximum entropy of integer-valued discrete random variables are obtained by applying the differential entropy results to continuous random variables which approximate the integer-valued random variables in a natural manner.*

*Most of these results are not new. The purpose of this article is to present all the results in an integrated framework that includes continuous and discrete random variables, constraints on the permissible value set, and all possible values of  $p$ . Understanding such as this is useful in evaluating the performance of data compression schemes.*

## I. Introduction

The *differential entropy*  $\mathbf{h}\{x\}$  of a continuous, real-valued random variable  $x$  with probability density  $f(x)$  is defined as

$$\mathbf{h}\{x\} = -\mathbf{E}\{\log[f(x)]\} = -\int_{-\infty}^{\infty} f(x) \log[f(x)] dx \quad (1)$$

For any positive (or infinite) integer  $p = 1, 2, 3, \dots, \infty$ , define the  $L_p$ -norm  $\mathbf{M}_p\{x\}$  of the random variable  $x$  as

$$\begin{aligned} \mathbf{M}_p\{x\} &= [\mathbf{E}\{|x|^p\}]^{1/p} \\ &= \left[ \int_{-\infty}^{\infty} f(x) |x|^p dx \right]^{1/p}, \quad p = 1, 2, 3, \dots \\ \mathbf{M}_\infty\{x\} &= \lim_{p \rightarrow \infty} \mathbf{M}_p\{x\} = \text{ess sup}_{f(x) > 0} |x| \end{aligned} \quad (2)$$

The essential supremum in Eq. (2) is the smallest number that upper bounds  $|x|$  almost surely.

Sometimes the real-valued random variable  $x$  is constrained to lie within a subset  $\Xi$  of the real line; in this case, the integrals in Eqs. (1) and (2) need only extend over the subset  $\Xi$ .

For a discrete random variable  $X$  with discrete value set  $\Xi = \{\xi_i\}$  and probability mass function  $F(\xi_i)$ , its (absolute) *entropy*  $\mathbf{H}\{X\}$  is defined as

$$\mathbf{H}\{X\} = -\mathbf{E}\{\log[F(X)]\} = -\sum_i F(\xi_i) \log[F(\xi_i)] \quad (3)$$

The  $L_p$ -norm  $\mathbf{M}_p\{X\}$  of the discrete random variable  $X$  is defined as

$$\begin{aligned} \mathbf{M}_p\{X\} &= [\mathbf{E}\{|X|^p\}]^{1/p} \\ &= \left[ \sum_i F(\xi_i) |\xi_i|^p \right]^{1/p}, \quad p = 1, 2, 3, \dots \\ \mathbf{M}_\infty\{X\} &= \lim_{p \rightarrow \infty} \mathbf{M}_p\{X\} = \sup_{F(\xi_i) > 0} |\xi_i| \end{aligned} \quad (4)$$

This article tabulates continuous probability density functions  $f(x) = f_p^*(x; \mu)$  or  $f(x) = f_p^*(x; \mu, \Xi)$  and discrete probability mass functions  $F(\xi_i) = F_p^*(\xi_i; \mu, \Xi)$  which maximize the differential entropy  $\mathbf{h}\{x\}$  or absolute

entropy  $\mathbf{H}\{X\}$ , respectively, among all probability distributions with a given  $L_p$ -norm  $\mathbf{M}_p\{x\}$  or  $\mathbf{M}_p\{X\}$  and unconstrained or constrained value set  $\Xi$ . The most interesting results are obtained and plotted for unconstrained continuous random variables and for integer-valued discrete random variables. Finally, alternative simpler bounds on the entropy of integer-valued random variables are obtained by modifying the bounds on differential entropy for unconstrained continuous random variables.

Most of these results are not new. In fact, the maximum-entropy continuous distributions for  $p = 1, 2$  (Laplacian and Gaussian distributions, respectively) have been known since Shannon's original work [1]. The purpose of this article is to present all the results in an integrated framework that includes continuous and discrete random variables, constraints on the permissible value set, and all possible values of  $p$ .

Throughout this article, regular italic notation is used for an ordinary function of a real variable, such as  $f(x)$  or  $F(\xi_i)$ , while boldface notation is used for an operator applied to a random variable, such as  $\mathbf{h}\{x\}$  or  $\mathbf{H}\{X\}$ ,  $\mathbf{M}_p\{x\}$  or  $\mathbf{M}_p\{X\}$ , or the expectation operator  $\mathbf{E}\{\cdot\}$ . In order not to interrupt the main presentation, proofs of all stated results are relegated to the Appendix.

## II. Effects of Elementary Transformations

A *scaled* random variable  $x' = qx$  or  $X' = qX$ , where  $q$  is a constant, has a correspondingly scaled  $L_p$ -norm:

$$\begin{aligned} \mathbf{M}_p\{x'\} &= |q| \mathbf{M}_p\{x\} \\ \mathbf{M}_p\{X'\} &= |q| \mathbf{M}_p\{X\} \end{aligned} \quad (5)$$

A discrete random variable  $X$  with value set  $\Xi = \{\xi_i\}$  scales to a discrete random variable  $X'$  with scaled value set  $q\Xi \equiv \{q\xi_i\}$ . The entropy of a discrete random variable is unaffected by scaling, but the differential entropy of a scaled continuous random variable either increases or decreases:

$$\begin{aligned} \mathbf{h}\{x'\} &= \mathbf{h}\{x\} + \log[|q|] \\ \mathbf{H}\{X'\} &= \mathbf{H}\{X\} \end{aligned} \quad (6)$$

The change in the differential entropy of a scaled continuous random variable exactly equals the change in the logarithm of its  $L_p$ -norm:

$$\mathbf{h}\{x'\} - \mathbf{h}\{x\} = \log[\mathbf{M}_p\{x'\}] - \log[\mathbf{M}_p\{x\}] = \log[|q|] \quad (7)$$

In contrast, the  $L_p$ -norm of a discrete random variable can be made arbitrarily small or large without affecting its entropy, simply by scaling its value set.

A *shifted* random variable  $x'' = x - \Delta$  or  $X'' = X - \Delta$ , where  $\Delta$  is a constant, has the same differential or absolute entropy as the unshifted random variable,

$$\begin{aligned} \mathbf{h}\{x''\} &= \mathbf{h}\{x\} \\ \mathbf{H}\{X''\} &= \mathbf{H}\{X\} \end{aligned} \quad (8)$$

but a different  $L_p$ -norm. A discrete random variable  $X$  with value set  $\Xi = \{\xi_i\}$  shifts to a discrete random variable  $X''$  with shifted value set  $\Xi - \Delta \equiv \{\xi_i - \Delta\}$ . A random variable  $x$  or  $X$  is *centered* with respect to the  $L_p$ -norm if no shifted version has a lower  $L_p$ -norm. A centered random variable  $x_p^o$  or  $X_p^o$  can be obtained from an uncentered random variable  $x$  or  $X$  by applying an optimum shift  $\Delta = \Delta_p^o$ . This optimum shift equals the median of the random variable for  $p = 1$ , the mean value of the random variable for  $p = 2$ , and the average of the essential infimum and essential supremum of the random variable for  $p = \infty$ . The centered  $L_p$ -norm  $\mathbf{M}_p^o\{x\}$  or  $\mathbf{M}_p^o\{X\}$  of the random variable  $x$  or  $X$  can be defined as

$$\begin{aligned} \mathbf{M}_p^o\{x\} &= \min_{\Delta} \mathbf{M}_p\{x - \Delta\} = \mathbf{M}_p\{x - \Delta_p^o\} = \mathbf{M}_p\{x_p^o\} \\ \mathbf{M}_p^o\{X\} &= \min_{\Delta} \mathbf{M}_p\{X - \Delta\} = \mathbf{M}_p\{X - \Delta_p^o\} = \mathbf{M}_p\{X_p^o\} \end{aligned} \quad (9)$$

### III. Maximum Differential Entropy for Continuous Random Variables

For any positive real number  $\mu$  and any positive (or infinite) integer  $p = 1, 2, \dots, \infty$ , let  $x_p^*(\mu)$  be a continuous random variable with probability density  $f_p^*(x; \mu)$ , where

$$\begin{aligned} f_p^*(x; \mu) &= \frac{\exp(-|x|^p/p\mu^p)}{2\mu p^{1/p} \Gamma\left(\frac{p+1}{p}\right)}, \quad p = 1, 2, 3, \dots \\ f_\infty^*(x; \mu) &= \begin{cases} \frac{1}{2\mu}, & |x| \leq \mu \\ 0, & |x| > \mu \end{cases} \end{aligned} \quad (10)$$

and  $\Gamma(\cdot)$  is the gamma function. These probability densities are all properly normalized, i.e.,

$$\int_{-\infty}^{\infty} f_p^*(x; \mu) dx = 1, \quad p = 1, 2, 3, \dots, \infty \quad (11)$$

The probability densities  $f_p^*(x; \mu)$  for  $p = 1, 2, \infty$  are the well-known Laplacian, Gaussian, and uniform probability densities, respectively.

The absolute moments of these random variables are known in closed form:

$$\begin{aligned} \mathbf{E}\{|x_p^*(\mu)|^n\} &= \mu^n \frac{\Gamma\left(\frac{n+1}{p}\right) p^{n/p}}{\Gamma\left(\frac{1}{p}\right)}, \\ n &= 1, 2, 3, \dots, \quad p = 1, 2, 3, \dots \\ \mathbf{E}\{|x_\infty^*(\mu)|^n\} &= \frac{\mu^n}{n+1}, \quad n = 1, 2, 3, \dots \end{aligned} \quad (12)$$

Evaluating these expressions for  $n = p$  or  $n \rightarrow \infty$  yields the  $L_p$ -norm  $M_p^*(\mu)$  of the random variable  $x_p^*(\mu)$ :

$$M_p^*(\mu) \equiv \mathbf{M}_p\{x_p^*(\mu)\} = \mu, \quad p = 1, 2, 3, \dots, \infty \quad (13)$$

The differential entropy  $h_p^*(\mu)$  of the random variable  $x_p^*(\mu)$  is calculated as

$$\begin{aligned} h_p^*(\mu) &\equiv \mathbf{h}\{x_p^*(\mu)\} \\ &= \log\left[2\mu \Gamma\left(\frac{p+1}{p}\right) (pe)^{1/p}\right], \quad p = 1, 2, 3, \dots \\ h_\infty^*(\mu) &\equiv \mathbf{h}\{x_\infty^*(\mu)\} = \log[2\mu] \end{aligned} \quad (14)$$

Explicit formulas for  $p = 1, 2$  are

$$\begin{aligned} h_1^*(\mu) &= \log[2e\mu] \\ h_2^*(\mu) &= \log[\sqrt{2\pi e} \mu] \end{aligned} \quad (15)$$

Since from Eq. (13) the parameter  $\mu$  equals the  $L_p$ -norm  $M_p^*(\mu)$  for any  $p$ , the differential entropy can be related directly to the corresponding  $L_p$ -norm:

$$\begin{aligned} h_p^*(\mu) &= \log\left[2 \Gamma\left(\frac{p+1}{p}\right) (pe)^{1/p}\right] + \log[M_p^*(\mu)], \\ p &= 1, 2, 3, \dots \\ h_\infty^*(\mu) &= \log[2] + \log[M_\infty^*(\mu)] \end{aligned} \quad (16)$$

The differential entropy  $h_p^*(\mu)$  is plotted in Fig. 1 versus the logarithm of the corresponding  $L_p$ -norm,  $\log[M_p^*(\mu)]$ , for various values of  $p$ . Note that this is a simple straight-line relationship. In fact, the straight line has unit slope, assuming  $\log[M_p^*(\mu)]$  is measured to the same logarithmic base as  $h_p^*(\mu)$ . This is consistent with the previous observation in Eq. (7), because the scaled version of the

random variable  $x_p^*(\mu)$  is statistically equivalent to the random variable with scaled  $L_p$ -norm, i.e.,

$$qx_p^*(\mu) \Leftrightarrow x_p^*(|q|\mu) \quad (17)$$

If  $x$  is any continuous random variable with differential entropy  $\mathbf{h}\{x\}$  and  $L_p$ -norm  $\mathbf{M}_p\{x\} = \mu$ , then

$$\mathbf{h}\{x\} \leq h_p^*(\mathbf{M}_p\{x\}) = \mathbf{h}\{x_p^*(\mu)\}, \quad p = 1, 2, 3, \dots, \infty \quad (18)$$

i.e.,  $x_p^*(\mu)$  is the *maximum-entropy* continuous random variable with a fixed  $L_p$ -norm  $\mu$ . Since the bound in Eq. (19) must be valid for all values of  $p$ ,

$$\mathbf{h}\{x\} \leq \min_p h_p^*(\mathbf{M}_p\{x\}) \quad (19)$$

If the random variable  $x$  is not centered with respect to the  $L_p$ -norm, the centered random variable  $x_p^o = x - \Delta_p^o$  has the same differential entropy as  $x$  but a smaller  $L_p$ -norm. The differential entropy of  $x$  may be more tightly upper bounded by applying the bounds in Eqs. (18) and (19) to the differential entropy of  $x_p^o$ :

$$\begin{aligned} \mathbf{h}\{x\} = \mathbf{h}\{x_p^o\} &\leq h_p^*(\mathbf{M}_p\{x_p^o\}) = h_p^*(\mathbf{M}_p^o\{x\}), \\ &p = 1, 2, 3, \dots, \infty \end{aligned} \quad (20)$$

and

$$\mathbf{h}\{x\} \leq \min_p h_p^*(\mathbf{M}_p^o\{x\}) \quad (21)$$

If the real-valued continuous random variable  $x$  is constrained to lie within a subset  $\Xi$  of the real line, its maximum possible differential entropy is smaller than that calculated above for a random variable constrained only by its  $L_p$ -norm. Maximum-entropy distributions for constrained continuous random variables can be obtained as simple generalizations of the foregoing results. Let  $x_p^*(\mu, \Xi)$  be a continuous random variable with probability density  $f_p^*(x; \mu, \Xi)$  equal to the conditional probability density of  $x_p^*(\mu)$  given  $\{x_p^*(\mu) \in \Xi\}$ , i.e.,

$$\begin{aligned} f_p^*(x; \mu, \Xi) &= \begin{cases} \frac{\exp(-|x|^p/p\mu^p)}{\alpha_p^*(\mu, \Xi)}, & x \in \Xi \\ 0, & x \notin \Xi \end{cases} \\ &p = 1, 2, 3, \dots \\ f_\infty^*(x; \mu, \Xi) &= \begin{cases} \frac{1}{\alpha_\infty^*(\mu, \Xi)}, & |x| \leq \mu \text{ and } x \in \Xi \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (22)$$

where

$$\begin{aligned} \alpha_p^*(\mu, \Xi) &= \int_{\Xi} \exp(-|x|^p/p\mu^p) dx, \quad p = 1, 2, 3, \dots \\ \alpha_\infty^*(\mu, \Xi) &= \int_{\Xi \cap \{|x| \leq \mu\}} 1 dx \end{aligned} \quad (23)$$

The  $L_p$ -norm  $M_p^*(\mu, \Xi)$  of the random variable  $x_p^*(\mu, \Xi)$  is given by

$$\begin{aligned} M_p^*(\mu, \Xi) &\equiv \mathbf{M}_p\{x_p^*(\mu, \Xi)\} \\ &= \mu \left[ \frac{\beta_p^*(\mu, \Xi)}{\alpha_p^*(\mu, \Xi)} \right]^{\frac{1}{p}}, \quad p = 1, 2, 3, \dots \\ M_\infty^*(\mu, \Xi) &\equiv \mathbf{M}_\infty\{x_\infty^*(\mu, \Xi)\} = \sup_{\substack{|x| \leq \mu \\ x \in \Xi}} |x| \end{aligned} \quad (24)$$

where

$$\begin{aligned} \beta_p^*(\mu, \Xi) &= \int_{\Xi} (|x|^p/\mu^p) \exp(-|x|^p/p\mu^p) dx, \\ &p = 1, 2, 3, \dots \end{aligned} \quad (25)$$

The differential entropy  $h_p^*(\mu, \Xi)$  of the random variable  $x_p^*(\mu, \Xi)$  is given by

$$\begin{aligned} h_p^*(\mu, \Xi) &\equiv \mathbf{h}\{x_p^*(\mu, \Xi)\} \\ &= \log[\alpha_p^*(\mu, \Xi)] + \frac{\log[e] \beta_p^*(\mu, \Xi)}{p \alpha_p^*(\mu, \Xi)} \\ &= \log[\alpha_p^*(\mu, \Xi)] + \frac{\log[e]}{p} \left[ \frac{M_p^*(\mu, \Xi)}{\mu} \right]^p, \\ &p = 1, 2, 3, \dots \end{aligned}$$

$$h_\infty^*(\mu, \Xi) \equiv \mathbf{h}\{x_\infty^*(\mu, \Xi)\} = \log[\alpha_\infty^*(\mu, \Xi)] \quad (26)$$

The random variable  $x_p^*(\mu, \Xi)$  is the *maximum-entropy* continuous random variable with constrained value set  $X$  and fixed  $L_p$ -norm  $M_p^*(\mu, \Xi)$ , i.e., if  $x$  is any continuous random variable with value set  $\Xi$ , differential entropy  $\mathbf{h}\{x\}$ , and  $L_p$ -norm  $\mathbf{M}_p\{x\}$ , then

$$\begin{aligned} \mathbf{h}\{x\} &\leq \mathbf{h}\{x_p^*(\mu_p, \Xi)\} \\ &= h_p^*(\mu_p, \Xi), \quad p = 1, 2, 3, \dots, \infty \end{aligned} \quad (27)$$

where  $\mu_p$  is chosen to match the  $L_p$ -norm of  $x$ :

$$M_p^*(\mu_p, \Xi) = \mathbf{M}_p\{x\}, \quad p = 1, 2, 3, \dots, \infty \quad (28)$$

Since the bound in Eq. (27) must be valid for all values of  $p$ ,

$$\mathbf{h}\{x\} \leq \min_p h_p^*(\mu_p, \Xi) \quad (29)$$

If the random variable  $x$  is not centered with respect to the  $L_p$ -norm, the differential entropy of  $x$  may be more tightly upper bounded by applying the bounds in Eqs. (27) and (29) to the differential entropy of the centered random variable  $x_p^o = x - \Delta_p^o$ :

$$\mathbf{h}\{x\} = \mathbf{h}\{x_p^o\} \leq \mathbf{h}\{x_p^*(\mu_p^o, \Xi - \Delta_p^o)\} = h_p^*(\mu_p^o, \Xi - \Delta_p^o), \quad p = 1, 2, 3, \dots, \infty \quad (30)$$

and

$$\mathbf{h}\{x\} \leq \min_p h_p^*(\mu_p^o, \Xi - \Delta_p^o) \quad (31)$$

where  $\mu_p^o$  is chosen to match the  $L_p$ -norm of  $x_p^o$  (i.e., the centered  $L_p$ -norm of  $x$ ):

$$\begin{aligned} M_p^*(\mu_p^o, \Xi - \Delta_p^o) &= \mathbf{M}_p\{x_p^o\} \\ &= \mathbf{M}_p^o\{x\}, \quad p = 1, 2, 3, \dots, \infty \end{aligned} \quad (32)$$

Notice that the bounds on the right-hand sides of Eqs. (30) and (31) are calculated with reference to the shifted value sets  $\Xi - \Delta_p^o$ , not the actual value set  $\Xi$ .

The integrals defining  $\alpha_p^*(\mu, \Xi)$  and  $\beta_p^*(\mu, \Xi)$  are generally not obtainable in closed form for an arbitrary value set  $\Xi$ . An interesting exception is when the value set equals the positive half-line, i.e.,  $\Xi = R^+ \equiv (0, \infty)$ . In this case,

$$M_p^*(\mu, R^+) = M_p^*(\mu) = \mu, \quad p = 1, 2, 3, \dots, \infty \quad (33)$$

and

$$h_p^*(\mu, R^+) = h_p^*(\mu) - \log[2], \quad p = 1, 2, 3, \dots, \infty \quad (34)$$

In other words, the maximum possible differential entropy for a positive-valued continuous random variable is exactly one bit less than the maximum differential entropy for a real-valued random variable with the same  $L_p$ -norm.

## IV. Maximum Entropy for Discrete Random Variables

Discrete versions  $F_p^*(\xi_i; \mu, \Xi)$  of the probability densities  $f_p^*(x; \mu)$  can be defined in a natural manner for discrete random variables  $X_p^*(\mu, \Xi)$  with discrete value set  $\Xi = \{\xi_i\}$ :

$$\begin{aligned} F_p^*(\xi_i; \mu, \Xi) &= \frac{f_p^*(\xi_i; \mu)}{\sum_j f_p^*(\xi_j; \mu)} \\ &= \frac{\exp(-|\xi_i|^p/p\mu^p)}{A_p^*(\mu, \Xi)}, \quad p = 1, 2, 3, \dots \end{aligned} \quad (35)$$

$$F_\infty^*(\xi_i; \mu, \Xi) = \frac{f_\infty^*(\xi_i; \mu)}{\sum_j f_\infty^*(\xi_j; \mu)} = \begin{cases} \frac{1}{A_\infty^*(\mu, \Xi)}, & |\xi_i| \leq \mu \\ 0, & |\xi_i| > \mu \end{cases}$$

where

$$\begin{aligned} A_p^*(\mu, \Xi) &= \sum_i \exp(-|\xi_i|^p/p\mu^p), \quad p = 1, 2, 3, \dots \\ A_\infty^*(\mu, \Xi) &= \sum_{|\xi_i| \leq \mu} 1 \end{aligned} \quad (36)$$

The discrete probability mass function  $F_p^*(\xi_i; \mu, \Xi)$  equals the conditional probability mass function for the maximum-entropy continuous random variable  $x_p^*(\mu)$ , given  $\{x_p^*(\mu) \in \Xi\}$ .

The  $L_p$ -norm  $M_p^*(\mu, \Xi)$  of the discrete random variable  $X_p^*(\mu, \Xi)$  is given by

$$\begin{aligned} M_p^*(\mu, \Xi) &\equiv \mathbf{M}_p\{X_p^*(\mu, \Xi)\} \\ &= \mu \left[ \frac{B_p^*(\mu, \Xi)}{A_p^*(\mu, \Xi)} \right]^{\frac{1}{p}}, \quad p = 1, 2, 3, \dots \\ M_\infty^*(\mu, \Xi) &\equiv \mathbf{M}_\infty\{X_\infty^*(\mu, \Xi)\} = \sup_{|\xi_i| \leq \mu} |\xi_i| \end{aligned} \quad (37)$$

where

$$\begin{aligned} B_p^*(\mu, \Xi) &= \sum_i (|\xi_i|^p/p\mu^p) \exp(-|\xi_i|^p/p\mu^p), \\ & \quad p = 1, 2, 3, \dots \end{aligned} \quad (38)$$

The entropy  $H_p^*(\mu, \Xi)$  of the discrete random variable  $X_p^*(\mu, \Xi)$  is given by

$$\begin{aligned} H_p^*(\mu, \Xi) &\equiv \mathbf{H}\{X_p^*(\mu, \Xi)\} \\ &= \log[A_p^*(\mu, \Xi)] + \frac{\log[e]}{p} \frac{B_p^*(\mu, \Xi)}{A_p^*(\mu, \Xi)} \\ &= \log[A_p^*(\mu, \Xi)] + \frac{\log[e]}{p} \left[ \frac{M_p^*(\mu, \Xi)}{\mu} \right]^p, \\ &\quad p = 1, 2, 3, \dots \\ H_\infty^*(\mu, \Xi) &\equiv \mathbf{H}\{X_\infty^*(\mu, \Xi)\} = \log[A_\infty^*(\mu, \Xi)] \quad (39) \end{aligned}$$

The random variable  $X_p^*(\mu, \Xi)$  is the *maximum-entropy* discrete random variable with value set  $\Xi$  and fixed  $L_p$ -norm  $M_p^*(\mu, \Xi)$ ; i.e., if  $X$  is any discrete random variable with value set  $\Xi$ , entropy  $\mathbf{H}\{X\}$ , and  $L_p$ -norm  $\mathbf{M}_p\{X\}$ , then

$$\begin{aligned} \mathbf{H}\{X\} &\leq \mathbf{H}\{X_p^*(\mu_p, \Xi)\} \\ &= H_p^*(\mu_p, \Xi), \quad p = 1, 2, 3, \dots, \infty \quad (40) \end{aligned}$$

where  $\mu_p$  is chosen to match the  $L_p$ -norm of  $X$ :

$$M_p^*(\mu_p, \Xi) = \mathbf{M}_p\{X\}, \quad p = 1, 2, 3, \dots, \infty \quad (41)$$

Since the bound in Eq. (40) must be valid for all values of  $p$ ,

$$\mathbf{H}\{X\} \leq \min_p H_p^*(\mu_p, \Xi) \quad (42)$$

If the random variable  $X$  is not centered with respect to the  $L_p$ -norm, the centered random variable  $X_p^o = X - \Delta_p^o$  has the same entropy as  $X$  but a smaller  $L_p$ -norm. The entropy of  $X$  may be more tightly upper bounded by applying the bounds in Eqs. (40) and (42) to the entropy of  $X_p^o$ .

$$\begin{aligned} \mathbf{H}\{X\} &= \mathbf{H}\{X_p^o\} \\ &\leq \mathbf{H}\{X_p^*(\mu_p^o, \Xi - \Delta_p^o)\} \\ &= H_p^*(\mu_p^o, \Xi - \Delta_p^o), \quad p = 1, 2, 3, \dots, \infty \quad (43) \end{aligned}$$

and

$$\mathbf{H}\{X\} \leq \min_p H_p^*(\mu_p^o, \Xi - \Delta_p^o) \quad (44)$$

where  $\mu_p^o$  is chosen to match the  $L_p$ -norm of  $X_p^o$  (i.e., the centered  $L_p$ -norm of  $X$ ):

$$\begin{aligned} M_p^*(\mu_p^o, \Xi - \Delta_p^o) &= \mathbf{M}_p\{X_p^o\} \\ &= \mathbf{M}_p^o\{X\}, \quad p = 1, 2, 3, \dots, \infty \quad (45) \end{aligned}$$

Notice again that the bounds based on centered random variables are calculated with reference to the shifted value sets  $\Xi - \Delta_p^o$ , not the actual value set  $\Xi$ . An exception for which the centering operation leaves the value set unchanged (i.e.,  $\Xi - \Delta_p^o = \Xi$ ) occurs for the value set  $\Xi = I$  (defined below) or, more generally, for any scaled version of it,  $\Xi = qI$ , as long as the allowable centering shifts  $\Delta_p^o$  are constrained to multiples of the scale quantum  $q$ .

For many applications, the most interesting discrete value sets are the set of all integers  $I \equiv \{0, \pm 1, \pm 2, \pm 3, \dots\}$  and the set of positive integers  $I^+ \equiv \{1, 2, 3, \dots\}$ . The maximum entropy for integer-valued random variables,  $H_p^*(\mu, I)$ , is plotted in Fig. 2 versus the logarithm of the corresponding  $L_p$ -norm,  $\log[M_p^*(\mu, I)]$ , for various values of  $p$ . Notice that the nonlinear relationship for integer-valued random variables becomes essentially linear when the  $L_p$ -norm is large compared to the (unit) interval between successive values in the value set  $I$ . In fact, all of the curves in Fig. 2 converge to the corresponding straight-line curves in Fig. 1 in the limit of large  $L_p$ -norm. Notice also how the continuous curves for large values of  $p < \infty$  approach the limiting staircase curve for  $p = \infty$ . The maximum entropy curve for  $p = \infty$  takes quantum jumps at integer values of the  $L_\infty$ -norm.

Closed-form maximum-entropy expressions as a function of  $L_p$ -norm can be obtained for discrete random variables in only a few special cases. Interesting cases include  $p = 1, \infty$ , for value sets  $\Xi = I, I^+$ :

$$\begin{aligned} H_1^*(\mu, I) &= \log \left[ M_1^*(\mu, I) + \sqrt{1 + [M_1^*(\mu, I)]^2} \right] \\ &\quad + M_1^*(\mu, I) \log \left[ \frac{M_1^*(\mu, I)}{\sqrt{1 + [M_1^*(\mu, I)]^2} - 1} \right] \end{aligned}$$

$$= M_1^*(\mu, I)$$

$$\begin{aligned} & \times \mathbf{H}_2 \left[ \frac{1 + M_1^*(\mu, I) - \sqrt{1 + [M_1^*(\mu, I)]^2}}{2M_1^*(\mu, I)} \right] \\ & + \frac{1 + M_1^*(\mu, I) + \sqrt{1 + [M_1^*(\mu, I)]^2}}{2} \\ & \times \mathbf{H}_2 \left[ \sqrt{1 + [M_1^*(\mu, I)]^2} - M_1^*(\mu, I) \right] \end{aligned}$$

$$H_\infty^*(\mu, I) = \log(2[M_\infty^*(\mu, I)] + 1) \quad (46)$$

and

$$\begin{aligned} H_1^*(\mu, I^+) &= M_1^*(\mu, I^+) \mathbf{H}_2 \left[ \frac{1}{M_1^*(\mu, I^+)} \right] \\ H_\infty^*(\mu, I^+) &= \log([M_\infty^*(\mu, I^+)]) \end{aligned} \quad (47)$$

where  $[a]$  is the integer part of  $a$  and  $\mathbf{H}_2[a]$  is the binary entropy function,

$$\mathbf{H}_2[a] = \begin{cases} -a \log[a] - (1-a) \log[1-a], & 0 < a < 1 \\ 0, & a = 0 \text{ or } a = 1 \end{cases} \quad (48)$$

## V. Alternative Entropy Bounds for Integer-Valued Random Variables

The maximum-entropy discrete distributions are not as useful as the maximum-entropy continuous distributions for unconstrained value sets, because closed-form results determining the maximum entropy for a given  $L_p$ -norm are available only in special cases. Alternative bounds on the entropy of discrete random variables can be obtained by approximating their discrete probability distributions with continuous probability densities and applying the simpler bounds on the differential entropy of continuous random variables. In this section, entropy bounds of this kind are obtained for integer-valued random variables ( $\Xi = I$ ).

Associate with any integer-valued random variable  $X$  a corresponding continuous random variable  $x$  defined by

$$x = X + u \quad (49)$$

where  $u$  is a uniform (continuous) random variable over  $[-1/2, 1/2]$  which is independent of  $X$ . The probability density function  $f(x)$  of the continuous random variable  $x$  is related to the probability mass function  $F(X)$  of the discrete random variable  $X$  as:

$$f(x) = F(\lfloor x + 1/2 \rfloor) \quad (50)$$

where  $\lfloor x + 1/2 \rfloor$  maps  $x$  to the nearest integer. The differential entropy of  $x$  equals the absolute entropy of  $X$ , i.e.,

$$\mathbf{h}\{x\} = \mathbf{H}\{X\} \quad (51)$$

and their  $L_p$ -norms are related as follows:

$$[\mathbf{M}_p\{x\}]^p = \begin{cases} \sum_{\substack{r=0 \\ r \text{ even}}}^{p-1} [\mathbf{M}_{p-r}\{X\}]^{p-r} \binom{p}{r} \frac{2^{-r}}{r+1} \\ + \frac{2^{-p}}{p+1} F(0), & p = 1, 3, 5, \dots \\ \sum_{\substack{r=0 \\ r \text{ even}}}^{p-2} [\mathbf{M}_{p-r}\{X\}]^{p-r} \binom{p}{r} \frac{2^{-r}}{r+1} \\ + \frac{2^{-p}}{p+1}, & p = 2, 4, 6, \dots \end{cases}$$

$$\mathbf{M}_\infty\{x\} = \mathbf{M}_\infty\{X\} + \frac{1}{2} \quad (52)$$

Explicit formulas for  $p = 1, 2, 3, 4$ , are:

$$\begin{aligned} \mathbf{M}_1\{x\} &= \mathbf{M}_1\{X\} + \frac{1}{4} F(0) \\ [\mathbf{M}_2\{x\}]^2 &= [\mathbf{M}_2\{X\}]^2 + \frac{1}{12} \\ [\mathbf{M}_3\{x\}]^3 &= [\mathbf{M}_3\{X\}]^3 + \frac{1}{4} \mathbf{M}_1\{X\} + \frac{1}{32} F(0) \\ [\mathbf{M}_4\{x\}]^4 &= [\mathbf{M}_4\{X\}]^4 + \frac{1}{2} [\mathbf{M}_2\{X\}]^2 + \frac{1}{80} \end{aligned} \quad (53)$$

The entropy of the integer-valued random variable  $X$  is upper bounded by

$$\mathbf{H}\{X\} = \mathbf{h}\{x\} \leq h_p^*(\mathbf{M}_p\{x\}), \quad p = 1, 2, 3, \dots, \infty \quad (54)$$

Explicit bounds for  $p = 1, 2, \infty$ , are:

$$\begin{aligned} \mathbf{H}\{X\} &\leq \log[2e] + \log \left[ \mathbf{M}_1\{X\} + \frac{1}{4}F(0) \right] \\ \mathbf{H}\{X\} &\leq \log[\sqrt{2\pi e}] + \frac{1}{2} \log \left[ [\mathbf{M}_2\{X\}]^2 + \frac{1}{12} \right] \\ \mathbf{H}\{X\} &\leq \log[2] + \log \left[ \mathbf{M}_\infty\{X\} + \frac{1}{2} \right] \end{aligned} \quad (55)$$

Since the bound in Eq. (54) is valid for all values of  $p$ ,

$$\mathbf{H}\{X\} \leq \min_p h_p^*(\mathbf{M}_p\{x\}) \quad (56)$$

The bound in Eq. (54) is not quite as tight as the achievable bound given earlier in Eq. (40), because the stepwise constant probability density of  $x = X + u$  given by Eq. (50) cannot exactly equal the maximum-entropy continuous probability density specified by Eq. (10). However, a stepwise-constant approximation can be very accurate when the probability distribution is much wider than the unit step width.

## VI. Summary and Potential Applications

This article has tabulated continuous probability density functions  $f(x) = f_p^*(x; \mu)$  or  $f(x) = f_p^*(x; \mu, \Xi)$  and discrete probability mass functions  $F(\xi_i) = F_p^*(\xi_i; \mu, \Xi)$  which maximize the differential entropy  $\mathbf{h}\{x\}$  or absolute entropy  $\mathbf{H}\{X\}$ , respectively, among all probability distributions with a given  $L_p$ -norm  $\mathbf{M}_p\{x\}$  or  $\mathbf{M}_p\{X\}$  and unconstrained or constrained value set  $\Xi$ . Expressions for the maximum entropy are evaluated as functions of the  $L_p$ -norm. These expressions are obtained in closed form for the case of unconstrained continuous random variables, and in this case there is a simple straight-line relationship between the maximum differential entropy and the logarithm of the  $L_p$ -norm. Corresponding expressions for discrete and constrained continuous random variables are given parametrically; closed-form expressions are available only for special cases. However, simpler alternative bounds on the maximum entropy of integer-valued random variables are obtained by applying the differential entropy results to continuous random variables which approximate the integer-valued random variables in a natural manner.

The results tabulated here have at least two potentially useful applications. First, they can lend a theoretical underpinning to source coding distortion measures based on  $L_p$ -norms. Second, they can be used to perform estimates of the local entropy of a dataset, for which the available local data are sufficient for obtaining good estimates of the dataset's  $L_p$ -norm but not for a good estimate of its histogram. Follow-up articles on these two applications will appear in future issues.

## Acknowledgments

Much of the impetus for the work in this article came from a derivation of the  $p = 2$  case of Eq. (55) in an unnumbered JPL memorandum by Bob McEliece. Mehrdad Shahshahani has pointed out that most of the results in this article remain valid for generalized  $L_p$ -norms with real-valued  $p$ , as long as  $p \geq 1$ . Bob McEliece has noted that similar maximum entropy expressions can be obtained when the  $L_p$ -norm constraint is replaced by a constraint on the expected value of an arbitrary function of the random variable.

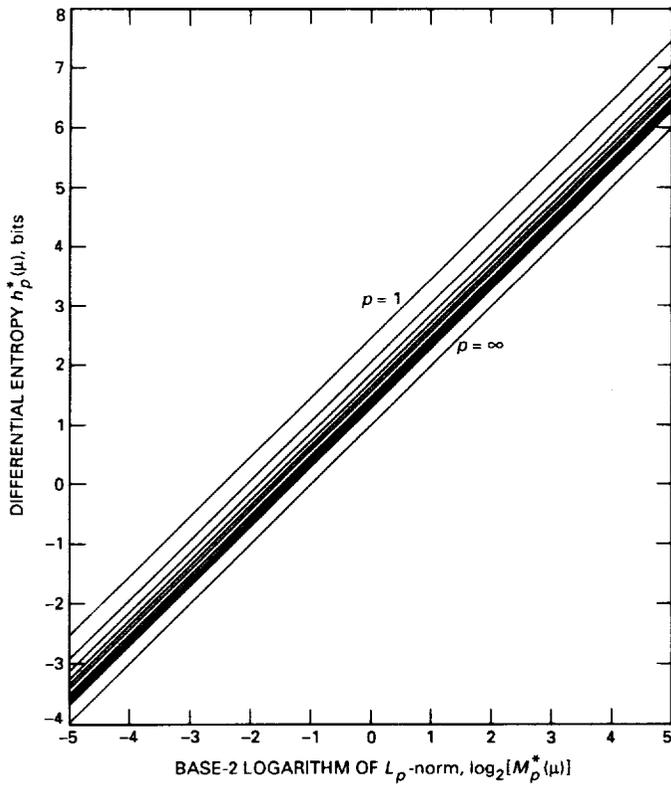


Fig. 1. Maximum differential entropy as a function of  $L_p$ -norm ( $p = 1, 2, 3, 4, 5, 6, 8, 10, 12, 16, \infty$ ) for unconstrained continuous random variables.

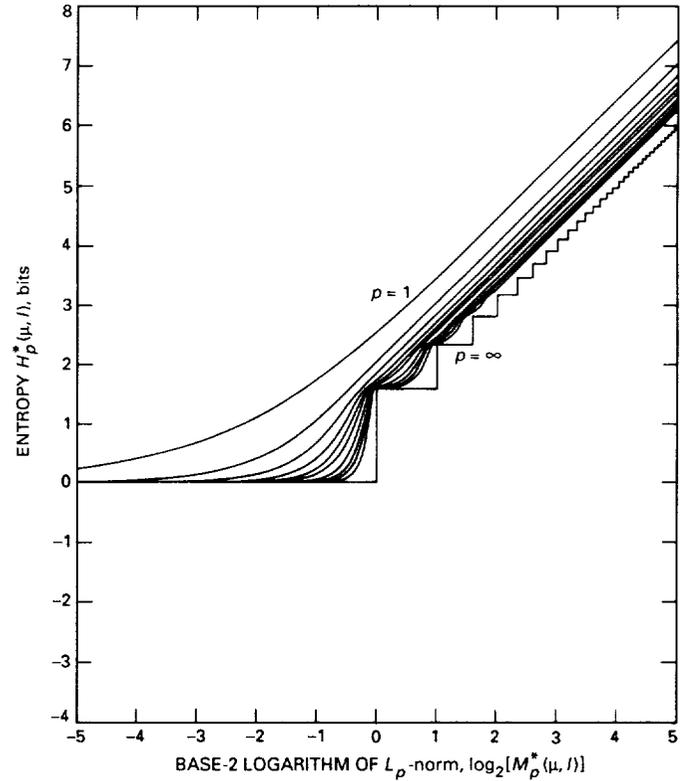


Fig. 2. Maximum entropy as a function of  $L_p$ -norm ( $p = 1, 2, 3, 4, 5, 6, 8, 10, 12, 16, \infty$ ) for integer-valued random variables.

## Appendix

This appendix contains proofs or derivations omitted in the main text. Equations (1), (2), (3), (4), (9), (10), (22), (23), (25), (28), (32), (35), (36), (38), (41), (45), (48), and (49) are definitions and require no proof. Equations (7), (16), (17), (19), (20), (21), (24), (29), (30), (31), (37), (42), (43), (44), (53), (54), (55), and (56) are trivial or straightforward applications of preceding results. This leaves Eqs. (5), (6), (8), (11), (12), (13), (14), (15), (18), (26), (27), (33), (34), (39), (40), (46), (47), (50), (51), and (52) requiring further justification.

Equation (5) follows from the linearity of the expectation operator. Equations (11) and (12) come from standard integral tables [2]. Equations (13) and (15) require two elementary properties [2] of the gamma function:  $\Gamma(1 + 1/p) = \Gamma(1/p)/p$  and  $\Gamma(1/2) = \sqrt{\pi}$ . Equations (6) and (8) result from applying the definitions in Eqs. (1) and (3) to the probability distributions of scaled and shifted random variables, obtained from standard texts [3] as:

$$\begin{aligned} f'(x') &= f(x'/q)/|q| & f''(x'') &= f(x'' + \Delta) \\ F'(X') &= F(X'/q) & F''(X'') &= F(X'' + \Delta) \end{aligned} \quad (\text{A-1})$$

where  $f'(x')$ ,  $F'(X')$ ,  $f''(x'')$ , and  $F''(X'')$  are probability density or probability mass functions for the scaled and shifted random variables  $x'$ ,  $X'$ ,  $x''$ , and  $X''$ .

Equations (14), (26), and (39) follow after observing that the logarithms of the probability distributions in Eqs. (10), (22), and (35) all consist of two terms, one term a constant and the second term proportional to  $|x|^p$  or  $|X|^p$ . The expected value of the second term can thus be calculated directly from the preceding formulas, Eqs. (13), (24), and (37), for the  $L_p$ -norm.

Equations (18), (27), and (40) are the central results of this article and are proved by generalizing a technique used in [4] to show that maximum differential entropy with constrained second moment is achieved by a Gaussian distribution. If  $x$  and  $x_p^*(\mu, \Xi)$  both have  $L_p$ -norm  $\mathbf{M}_p\{x\}$ , then for  $p < \infty$ ,

$$\begin{aligned} \mathbf{h}\{x_p^*(\mu, \Xi)\} &= - \int_{\Xi} f_p^*(x; \mu, \Xi) \log[f_p^*(x; \mu, \Xi)] dx \\ &= \int_{\Xi} f_p^*(x; \mu, \Xi) \left\{ \log[\alpha_p^*(\mu, \Xi)] + \frac{\log[e]}{p} \frac{|x|^p}{\mu^p} \right\} dx \\ &= \int_{\Xi} f(x) \left\{ \log[\alpha_p^*(\mu, \Xi)] + \frac{\log[e]}{p} \frac{|x|^p}{\mu^p} \right\} dx \\ &= - \int_{\Xi} f(x) \log[f_p^*(x; \mu, \Xi)] dx \end{aligned} \quad (\text{A-2})$$

The third equality in Eq. (A-2) follows from the assumption that  $x$  and  $x_p^*(\mu, \Xi)$  have identical  $L_p$ -norms, hence  $|x|^p$  has the same expectation whether it is averaged over  $f_p^*(x; \mu, \Xi)$  or  $f(x)$ . If  $p = \infty$ , the same result holds: the second term in the second and third lines of Eq. (A-2) is absent, and the integration over  $\Xi$  is replaced by an integration over  $\Xi \cap \{|x| \leq \mu\}$ . Continuing,

$$\begin{aligned} \mathbf{h}\{x_p^*(\mu, \Xi)\} - \mathbf{h}\{x\} &= - \int_{\Xi} f(x) \log[f_p^*(x; \mu, \Xi)] dx + \int_{\Xi \cap \{f(x) > 0\}} f(x) \log[f(x)] dx \\ &= \int_{\Xi \cap \{f(x) > 0\}} f(x) \log \left[ \frac{f(x)}{f_p^*(x; \mu, \Xi)} \right] dx \\ &\geq \int_{\Xi \cap \{f(x) > 0\}} f(x) \log[c] \left\{ 1 - \frac{f_p^*(x; \mu, \Xi)}{f(x)} \right\} dx = 0 \end{aligned} \quad (\text{A-3})$$

The inequality in Eq. (A-3) results from the general inequality  $\log[a] \geq \log[e](1 - 1/a)$  for all  $a > 0$ , and the last equality arises because  $f_p^*(x; \mu, \Xi)$  and  $f(x)$  both integrate to one.

The derivation in Eqs. (A-2) and (A-3) proves Eq. (27). Equation (18) is a special case of Eq. (27) obtained by setting  $\Xi$  equal to the set of all real numbers. Equation (40) is derived in a similar manner by replacing the integrals in Eqs. (A-2) and (A-3) with summations and continuous probability density functions with discrete probability mass functions.

Equation (33) results from noting that  $|x_p^*(\mu, R^+)| \Leftrightarrow |x_p^*(\mu)|$ , so the  $L_p$ -norms of  $x_p^*(\mu, R^+)$  and  $x_p^*(\mu)$  must be identical. Equation (34) comes from the fact that the constant scale factor  $\alpha_p^*(\mu, R^+)$  for  $f_p^*(x; \mu, R^+)$  in Eq. (22) with  $\Xi = R^+$  is exactly half the corresponding scale factor for  $f_p^*(x; \mu)$  in Eq. (10). This accounts for a difference of  $\log[2]$  in the first terms in their respective expressions for differential entropy. The second terms must be equal by the previous observation linking them to their respective  $L_p$ -norms.

To derive Eqs. (46) and (47), let  $a = e^{-1/\mu}$  and replace  $\Xi$  with  $I$  or  $I^+$  in Eqs. (36), (37), and (38) to obtain

$$A_1^*(\mu, I) = \sum_{i=-\infty}^{\infty} e^{-|i|/\mu} = 2 \sum_{i=0}^{\infty} a^i - 1 = \frac{2}{1-a} - 1 = \frac{1+a}{1-a}$$

$$\mu B_1^*(\mu, I) = \sum_{i=-\infty}^{\infty} |i| e^{-|i|/\mu} = 2 \sum_{i=0}^{\infty} i a^i = \frac{2a}{(1-a)^2}$$

$$A_{\infty}^*(\mu, I) = \sum_{|i| \leq \mu} 1 = 2\lfloor \mu \rfloor + 1$$

$$M_{\infty}^*(\mu, I) = \sup_{|i| \leq \mu} |i| = \lfloor \mu \rfloor \tag{A-4}$$

and

$$A_1^*(\mu, I^+) = \sum_{i=1}^{\infty} e^{-|i|/\mu} = \sum_{i=1}^{\infty} a^i = \frac{1}{1-a} - 1 = \frac{a}{1-a}$$

$$\mu B_1^*(\mu, I^+) = \sum_{i=1}^{\infty} |i| e^{-|i|/\mu} = \sum_{i=1}^{\infty} i a^i = \frac{a}{(1-a)^2}$$

$$A_{\infty}^*(\mu, I^+) = \sum_{1 \leq i \leq \mu} 1 = \lfloor \mu \rfloor$$

$$M_{\infty}^*(\mu, I^+) = \sup_{1 \leq i \leq \mu} |i| = \lfloor \mu \rfloor \tag{A-5}$$

where  $\lfloor \mu \rfloor$  is the integer part of  $\mu$ . The entropy expressions in Eqs. (46) and (47) follow algebraically upon substitution of Eqs. (A-4) and (A-5) into Eqs. (37) and (39) and solving for the entropy in terms of the corresponding  $L_p$ -norm.

Equation (50) results from calculating the conditional probability density  $f(x|X)$  of  $x$  given  $X$ , then averaging over  $X$ :

$$f(x|X) = \begin{cases} 1, & \text{if } |x - X| \leq 1/2 \\ 0, & \text{if } |x - X| > 1/2 \end{cases} = \begin{cases} 1, & \text{if } X = \lfloor x + 1/2 \rfloor \\ 0, & \text{if } X \neq \lfloor x + 1/2 \rfloor \end{cases}$$

$$f(x) = \sum_{i=-\infty}^{\infty} F(i)f(x|X=i) = F(\lfloor x + 1/2 \rfloor) \quad (\text{A-6})$$

Equation (51) results from breaking up the defining integral in Eq. (1) for the differential entropy into a sum of integrals over unit intervals,

$$\begin{aligned} \mathbf{h}\{x\} &= - \int_{-\infty}^{\infty} f(x) \log[f(x)] dx = - \sum_{i=-\infty}^{\infty} \int_{i-1/2}^{i+1/2} f(x) \log[f(x)] dx \\ &= - \sum_{i=-\infty}^{\infty} \int_{i-1/2}^{i+1/2} F(i) \log[F(i)] dx = - \sum_{i=-\infty}^{\infty} F(i) \log[F(i)] = \mathbf{H}\{X\} \end{aligned} \quad (\text{A-7})$$

Equation (52) is derived by considering the cases of even and odd values of  $p$  separately. In the first case, when  $p$  is even,

$$\begin{aligned} \mathbf{E}\{|X + u|^p\} &= \mathbf{E}\{(X + u)^p\} = \sum_{r=0}^p \binom{p}{r} \mathbf{E}\{X^{p-r}\} \mathbf{E}\{u^r\} \\ &= \sum_{\substack{r=0 \\ r \text{ even}}}^p \binom{p}{r} \mathbf{E}\{X^{p-r}\} \frac{2^{-r}}{r+1} \end{aligned} \quad (\text{A-8})$$

because

$$\mathbf{E}\{u^r\} = \begin{cases} \frac{2^{-r}}{r+1}, & \text{if } r \text{ is even} \\ 0, & \text{if } r \text{ is odd} \end{cases} \quad (\text{A-9})$$

Thus, since  $X^{p-r} = |X|^{p-r}$  when  $p$  and  $r$  are both even,

$$\mathbf{E}\{|X + u|^p\} = \frac{2^{-p}}{p+1} + \sum_{\substack{r=0 \\ r \text{ even}}}^{p-2} \binom{p}{r} \frac{2^{-r}}{r+1} \mathbf{E}\{|X|^{p-r}\} \quad (\text{A-10})$$

In the second case, when  $p$  is odd, the derivation begins by writing

$$|X + u| = |X| + w \quad (\text{A-11})$$

where

$$w = \begin{cases} +u, & X \geq 1 \\ |u|, & X = 0 \\ -u, & X \leq -1 \end{cases} \quad (\text{A-12})$$

This decomposition is valid because  $X$  is integer valued. The conditional moments of  $w$  are

$$\mathbf{E}\{w^r|X\} = \begin{cases} \frac{2^{-r}}{r+1}, & \text{if } r \text{ is even or if } r \text{ is odd and } X = 0 \\ 0, & \text{if } r \text{ is odd and } X \neq 0 \end{cases} \quad (\text{A-13})$$

Thus,

$$\begin{aligned} \mathbf{E}\{|X + u|^p\} &= \mathbf{E}\{(|X| + w)^p\} = \mathbf{E}\left\{\sum_{r=0}^p \binom{p}{r} |X|^{p-r} w^r\right\} \\ &= \sum_{i \neq 0} F(i) \sum_{r=0}^p \binom{p}{r} |i|^{p-r} \mathbf{E}\{w^r|X=i\} + F(0) \mathbf{E}\{w^p|X=0\} \\ &= \sum_{i \neq 0} F(i) \sum_{\substack{r=0 \\ r \text{ even}}}^p \binom{p}{r} |i|^{p-r} \frac{2^{-r}}{r+1} + F(0) \frac{2^{-p}}{p+1} \\ &= \sum_{\substack{r=0 \\ r \text{ even}}}^{p-1} \binom{p}{r} \frac{2^{-r}}{r+1} \sum_{i \neq 0} F(i) |i|^{p-r} + F(0) \frac{2^{-p}}{p+1} \\ &= \sum_{\substack{r=0 \\ r \text{ even}}}^{p-1} \binom{p}{r} \frac{2^{-r}}{r+1} \mathbf{E}\{|X|^{p-r}\} + F(0) \frac{2^{-p}}{p+1} \end{aligned} \quad (\text{A-14})$$

## References

- [1] C. E. Shannon "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 (Part I), pp. 623–656 (Part II), 1948; reprinted in book form with postscript by W. Weaver, Urbana: University of Illinois Press, 1949.
- [2] *CRC Standard Mathematical Tables*, ed. W. H. Beyer, Boca Raton, Florida: CRC Press, Inc., 1987.
- [3] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill, 1965.
- [4] R. E. Blahut, *Principles and Practice of Information Theory*, Reading, Massachusetts: Addison-Wesley, 1988.

58-90

332666

P-10

TDA Progress Report 42-104

February 15, 1991

N91-18316

JJB/450

## Compressed/Reconstructed Test Images for CRAF/Cassini

S. Dolinar, K.-M. Cheung, I. Onyszchuk, F. Pollara, and S. Arnold  
Communications Systems Research Section

*This article briefly describes a set of compressed, then reconstructed, test images submitted to the CRAF/Cassini project as part of its evaluation of near-lossless high-compression algorithms for representing image data. A total of seven test image files were provided by the project.*

*The seven test images have been compressed, then reconstructed with high quality (root-mean-square error of approximately one or two gray levels on an 8-bit gray scale), using discrete cosine transforms or Hadamard transforms and efficient entropy coders. The resulting compression ratios varied from about 2:1 to about 10:1, depending on the activity or randomness in the source image. This was accomplished without any special effort to optimize the quantizer or to introduce special postprocessing to filter the reconstruction errors.*

*A more complete set of measurements, showing the relative performance of the compression algorithms over a wider range of compression ratios and reconstruction errors, shows that additional compression is possible at a small sacrifice in fidelity.*

### I. Introduction

This article briefly describes a set of compressed, then reconstructed, test images submitted to the CRAF/Cassini project as part of its evaluation of near-lossless high-compression algorithms for representing image data. A total of seven test image files were provided by the project. Five test images (d1, f2, h2, j1, and l2) are star fields from the Hubble Space Telescope, and two images (saturn1 and saturn2) are views of Saturn from Voyager. Three of these original images are shown in Figs. 1(a), 1(b), and 1(c). The dimensions of the Hubble images and

the Saturn images are  $256 \times 256$  and  $800 \times 800$ , respectively. All images are represented by 8-bit pixel values in the range 0 to 255.

A total of 12 compressed/reconstructed images were returned to the project, as listed in Table 1. Three of the reconstructed images (marked by arrows  $\Rightarrow$  in Table 1) are shown in Figs. 2(a), 2(b), and 2(c) for comparison with the originals. Each of the seven test images was compressed using an algorithm that produces high quality reconstructed images (left-hand portion of Table 1)

with a root-mean-square error (RMSE) of about one gray level. Alternate compressed/reconstructed image versions (right-hand portion of Table 1) were also provided for five of the seven test images. Three of these alternate images show how much additional compression is possible at a small sacrifice in image fidelity, and the other two alternate images illustrate the effectiveness of a different compression algorithm, which is simpler to implement on the spacecraft.

In Table 1, the RMSE is computed in absolute units on an 8-bit gray level scale. The quoted bit rate is the number of bits per pixel required to encode the compressed image before reconstruction. The compression ratio is calculated as 8 bits divided by the bit rate.

## II. Description of the Compression System

The specific algorithms used to compress, then reconstruct the twelve images listed in Table 1 can be described with reference to the block diagram in Fig. 3. The various blocks in this diagram are described in the following sections.

### A. Data Transform/Inverse Data Transform

A discrete cosine transform (DCT) was applied to obtain ten of the twelve compressed/reconstructed images. The DCT is near-optimal for a wide variety of images, and is fast becoming an industry standard for high compression. The DCT was calculated using floating point arithmetic and applied to  $8 \times 8$  sub-blocks of the image.

A Hadamard transform (HT), also applied to  $8 \times 8$  blocks, was used for the remaining two compressed/reconstructed images. The HT is generally not as effective as the DCT, but it performed reasonably well for the seven test images. The HT is simpler to implement than the DCT, because it can be computed with integer arithmetic and without multiplications.

Mathematically, both transforms are defined as unitary transformations on each  $8 \times 8$  block of data. The image array  $X$  is decomposed into  $8 \times 8$  blocks  $X^{IJ}$ , i.e.,  $X = [X^{IJ}]$ , and the array of transform coefficients  $T$  is built from  $8 \times 8$  blocks  $T^{IJ}$ , i.e.,  $T = [T^{IJ}]$ . The blocks of transform coefficients are given by

$$T^{IJ} = \begin{cases} \frac{1}{8} CX^{IJ}C^T & \text{for DCT} \\ \frac{1}{8} HX^{IJ}H & \text{for HT} \end{cases}$$

where the  $8 \times 8$  matrices  $C = [c_{ij}]$  and  $H = [h_{ij}]$  are defined by

$$c_{ij} = \begin{cases} \sqrt{2} \cos\left(\frac{\pi i(2j+1)}{16}\right) & i = 1, 2, 3, 4, 5, 6, 7 \\ & j = 0, 1, 2, 3, 4, 5, 6, 7 \\ 1 & i = 0, j = 0, 1, 2, 3, 4, 5, 6, 7 \end{cases}$$

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix} = H^T$$

Both the DCT and the HT are exactly invertible in principle. The inverse transform formulas are simply:

$$X^{IJ} = \begin{cases} \frac{1}{8} C^T T^{IJ} C & \text{for DCT} \\ \frac{1}{8} H^T T^{IJ} H & \text{for HT} \end{cases}$$

Both the DCT and HT can be implemented with fast algorithms requiring fewer arithmetic operations than direct implementation of the matrix multiplications in the above expressions. The DCT requires real multiplications and additions, whereas the HT requires only integer additions and no multiplications. For this study, the DCT's real arithmetic was approximated by 32-bit floating point multiplications and additions, whereas integer arithmetic must generally be substituted in practice. Integer approximations in the computation of the transform coefficients can produce additional errors in the reconstructed images.

Transforming a block of data does not change its information content. Useful transforms concentrate most of the data's energy into a small number of transform coefficients. Low-energy transform coefficients can be encoded with a small number of bits.

### B. Quantizer/Dequantizer

The DCT produces real transform coefficients, which of necessity must be quantized to a finite number of bits. The HT produces quantized coefficients, but the quantization is impractically fine: the transform coefficients have eight times the dynamic range and one-eighth the granularity of the input, requiring six extra bits to represent exactly. So additional quantization is also performed for the HT.

Once the transform coefficients have been quantized, the dequantizer in Fig. 3 can only reconstruct an approximate version of the true coefficients, and the inverse data transform can no longer regenerate the exact original image. Except for the possible errors (noted above) in computing the transform coefficients, quantization of the computed coefficients is the only step in Fig. 3 that introduces errors in the reconstructed image data. The choice of quantization coarseness and uniformity thus sets the fidelity of the reconstructed image. This choice also limits the extent to which the entropy encoder can compress the image.

A uniform quantizer was used for all twelve of the compressed/reconstructed images. Mathematically, the output of the quantizer is an array of  $8 \times 8$  blocks,  $Q = [Q^{IJ}]$ , where each block,  $Q^{IJ} = [Q_{ij}^{IJ}]$ , is obtained from the  $8 \times 8$  block of transform coefficients,  $T^{IJ} = [T_{ij}^{IJ}]$ , as

$$Q_{ij}^{IJ} = \left\lfloor \frac{T_{ij}^{IJ}}{q} + \frac{1}{2} \right\rfloor$$

where  $q$  is the quantizer step size and  $\lfloor a \rfloor$  is the largest integer less than or equal to  $a$ . This uniform quantizer is actually “triply uniform.” Not only are the quantization levels for each transform coefficient equally spaced, but the quantization step size  $q$  is the same for all 64 coefficients within each  $8 \times 8$  block, and the step size does not change from block to block. A step size of  $q = 4$  was used for nine of the twelve images, and a coarser step size of  $q = 8$  was used for three of the alternate images.

Nonuniform quantization rules are available [1,2] to match the quantizer to the human visual response by selectively quantizing low-frequency DCT or HT coefficients more finely than high-frequency coefficients. Similarly, other algorithms can adapt the quantizer to the local statistics of the data on a block by block basis. However, such quantizers have so many adjustable parameters that a nonuniform quantizer optimized for a small set of test images would not fairly reflect the performance of the quantizer for untested images. Research is ongoing to find universal nonuniform, adaptive quantizers that consistently outperform the uniform quantizer.

### C. Entropy Encoder/Entropy Decoder

The entropy encoder losslessly encodes the array of quantized transform coefficients  $Q$  into a bit stream  $\mathbf{b}$  with bit rate approaching the entropy per coefficient of  $Q$ . Several types of highly efficient coders are available for this purpose. Among these are the Gallager–van-Voorhis–Huffman (GVH) coder [3], a variant of IBM’s arithmetic

Q-coder [4] being developed for the Joint Photographic Experts Group (JPEG) standard [2], and a baseline Huffman coder for this same standard. Variations of these coders were used to compress, then reconstruct the images listed in Table 1.

The first step in all three coding schemes is to arrange the quantized transform coefficients  $Q^{IJ} = [Q_{ij}^{IJ}]$  into an ordered sequence, starting with the DC coefficient ( $i = j = 0$ ) in the upper left-hand corner of the transformed block. The remaining 63 coefficients (AC coefficients) are ordered in some fashion, generally via a zigzag readout starting at the upper left-hand corner and working toward the lower right-hand corner. This zigzag sequence arranges the AC coefficients in increasing order of spatial frequency.

The GVH coding scheme is derived based on two observations on the quantized AC and DC coefficients. First, the AC coefficients and the differences between adjacent DC coefficients have two-sided geometric distributions. Second, runs of zeros occur frequently in the zigzag sequences of AC coefficients, especially at high compression ratios. By extending a result originally shown by Gallager and van Voorhis, a near-optimal adaptive coding scheme for prefix coding the two-sided geometric source is derived, avoiding both binning calculations and the Huffman tree generation algorithm. Instead, this scheme estimates the local activity of each  $8 \times 8$  block by counting the number of zeros in the block or in some preceding blocks, and adaptively encodes the transform coefficients using simple pipelined table lookup operations. An optional runlength code can also be used to encode runs of zeros in the zigzag sequence of AC coefficients.

The Q-coder is a lossless, binary entropy coder, developed by researchers at IBM, that efficiently implements an Elias code [5] on an input bit sequence. A coarsely quantized approximation to the real interval  $[0,1]$ , or a scaled version thereof, is recursively subdivided into two sections, whose sizes are proportional to probability estimates that the bit currently being coded is a 0 or 1. By dynamically updating these estimates using a finite-state machine, the Q-coder adapts to input data statistics (unlike a Huffman encoder, which requires statistics before coding), which makes it both robust and efficient. A coding model forms a binary sequence from a raster-scan ordering of the 64 quantized DCT or HT coefficients in each block. For each integer, an equal-to-zero flag bit, the sign bit, the position of the most significant bit, and then the least significant bits are sent through the Q-coder. Runlength coding was not performed because the small quantizer step sizes used make it unprofitable. A simple model, using only 12 prob-

ability estimates, each of which is a 5-bit number (state), was used instead of the complex model in Section 8 of [2] that requires 252 estimates. Negligible bit rate reductions are expected with the latter model for the images tested. The output sequence length is very close to a value calculated from the input stream entropy. Since only additions, subtractions, and comparisons are utilized by a Q-coder, it is simple and fast in practice. The particular variant implemented is described in Section 12 of [2].

The Joint Photographic Experts Group of the International Standards Organization/International Telegraph and Telephone Consultative Committee (ISO/CCITT) [2] is currently developing an international standard for still-image compression. In its baseline version, the proposed algorithm consists of an  $8 \times 8$  DCT, coefficient quantization, and Huffman or arithmetic coding. This scheme provides a near-lossless, high-compression image coding capability, which preserves image fidelity at compression rates competitive or superior to most known techniques. The DCT's 64 coefficients are independently uniformly quantized with a different step size for each coefficient. The DC component is differentially encoded, and the AC components are runlength encoded. Finally, some of the most significant bits of each resulting code are further encoded with a variable length code; the remaining bits are transmitted as they are. In the case of Huffman encoding, the JPEG default tables were used. The tables for the Huffman codes can be easily customized to adapt to the particular image source of interest.

The bit rates listed in Table 1 are the bit rates achieved by a variant of the Q-coder. The GVH and JPEG coders achieve comparable bit rates averaging 0.2 to 0.3 bits per pixel higher than the Q-coder for the images in Table 1.

#### D. Noiseless Channel

Because the entropy code is lossless, the entropy decoder is able to reconstruct an exact replica of the quantized transform coefficients, given the compressed files of coded bits. However, if the channel in Fig. 3 were not noiseless, the decoding process would be severely disrupted and errors might propagate wildly.

#### E. Preprocessing and Postprocessing (Not Implemented)

The original image data supplied by the project were actually obtained by preprocessing 12-bit data available from the cameras. The 12-bit data were subjected to a square-root operation, then quantized to 8 bits. No additional preprocessing was performed on the 8-bit data before the transform operation depicted in Fig. 3. The 8-bit

data provided by the project were considered to be the original image data for the purposes of the tests reported here.

Some sort of postprocessing is often desirable, following the inverse data transform, to make the reconstruction errors less noticeable. Postprocessing can be applied to remove visually disturbing blockiness in images that have been highly compressed by block-transform techniques. No such postprocessing was performed for any of the twelve compressed/reconstructed images, because it would unfairly mask the true efficacy of the compression algorithms.

A crude form of nonoptimum postprocessing actually did take place after the inverse transform, because the output values had to be quantized to 8 bits to fit the original image format. This quantization step should be skipped or deferred if the output data are subjected to further postprocessing (such as removal of the square-root operation mentioned above).

### III. Additional Performance Results

The twelve reconstructed images listed in Table 1, including the three images shown in Figs. 2(a), 2(b), and 2(c), were chosen to reflect the desires of the project to obtain compression ratios in the range of about 2:1 to about 10:1 with essentially zero reconstruction error. A more complete set of measurements showing the relative performance of the compression algorithms over a wider range of compression ratios and reconstruction errors was also obtained. These results are plotted in Figs. 4(a), 4(b), 4(c), 5(a), 5(b), and 5(c) for the three images shown in Figs. 1(a), 1(b), and 1(c), using DCT-based algorithms. Figures 4(a), 4(b), and 4(c) show the bit rate (bits per pixel) of the compressed images as a function of the RMSE distortion, and Figs. 5(a), 5(b), and 5(c) show the corresponding compression ratios. In these figures the bit rates and compression ratios achieved by the three entropy coders are compared with each other and with an approximate bound based on the estimated entropy of the quantized DCT coefficients. This entropy "bound" is derived assuming stationary statistics throughout the image; it can sometimes be beaten by algorithms capable of adapting to locally varying statistics.

### IV. Summary

The seven test images have been compressed, then reconstructed with high quality (RMSE of approximately one or two gray levels on an 8-bit gray scale) using DCT- or HT-based schemes and efficient entropy coders. The re-

sulting compression ratios varied from about 2:1 to about 10:1, depending on the activity or randomness in the source image. This was accomplished without making any special effort to optimize the quantizer or to introduce special postprocessing to filter the reconstruction errors.

A more complete set of measurements, showing the relative performance of the compression algorithms over a wider range of compression ratios and reconstruction errors, shows that additional compression is possible at a small sacrifice in fidelity.

## References

- [1] J. D. Eggerton and M. D. Srinath, "A Visually Weighted Quantization Scheme for Image Bandwidth Compression at Low Data Rates," *IEEE Trans. Commun.*, vol. COM-34, no. 8, pp. 840-847, August 1986.
- [2] *Joint Photographic Experts Group (JPEG) Draft Technical Specification (Rev. 5)*, ISO/CCITT, Washington, D.C., January 15, 1990.
- [3] K.-M. Cheung, P. Smyth, and H. Wang, "A High-Speed Distortionless Predictive Image-Compression Scheme," *TDA Progress Report 42-102*, vol. April-June 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 73-90, August 15, 1990.
- [4] W. B. Pennebaker, J. L. Mitchell, G. G. Langdon, Jr., and R. B. Arps, "An Overview of the Basic Principles of the Q-Coder Adaptive Binary Arithmetic Coder," *IBM J. Res. Develop.*, vol. 32, no. 6, pp. 717-726, November 1988.
- [5] R. E. Blahut, *Digital Transmission of Information*, Reading, Massachusetts: Addison-Wesley, p. 308, 1990.

**Table 1. List of original and compressed/reconstructed test images. Arrows  $\Rightarrow$  denote the three images shown in Figs. 2(a), 2(b), and 2(c).**

Original images	Compressed/Reconstructed images <sup>a</sup>			Alternate compressed/reconstructed images <sup>b</sup>		
	RMSE, absolute	Bit rate, bits/pixel	Compression ratio	RMSE, absolute	Bit rate, bits/pixel	Compression ratio
d1	1.19	1.50	5.33	—	—	—
f2	1.15	1.15	6.96	—	—	—
h2	1.16	1.21	6.61	(H) 1.18	1.25	6.40
j1	1.19	3.77	2.12	(D) 2.33	2.56	3.12
l2	1.19	2.04	3.92	$\Rightarrow$ (D) 2.25	1.10	7.27
saturn1	0.97	1.31	6.11	$\Rightarrow$ (D) 1.43	0.75	10.67
saturn2	$\Rightarrow$ 0.87	0.82	9.76	(H) 0.95	1.00	8.00

<sup>a</sup> Using Q-coder and DCT with quantization step size  $q = 4$ .

<sup>b</sup> Using Q-coder and either DCT with quantization step size  $q = 8$  (D) or HT with step size  $q = 4$  (H).

ORIGINAL PAGE  
BLACK AND WHITE PHOTOGRAPH

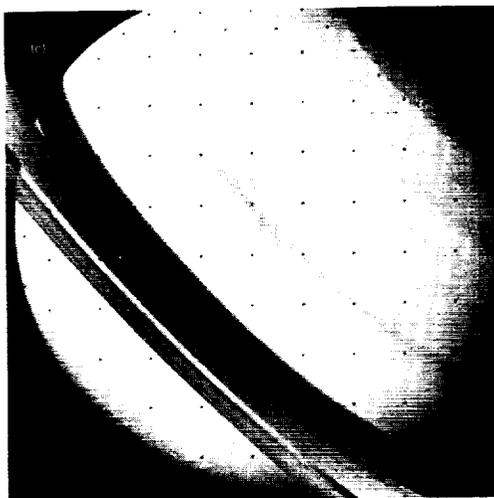
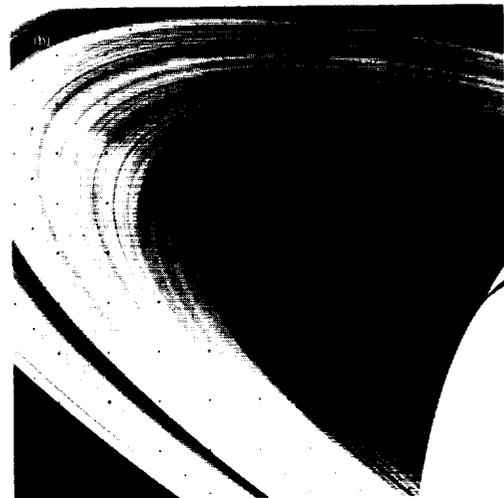
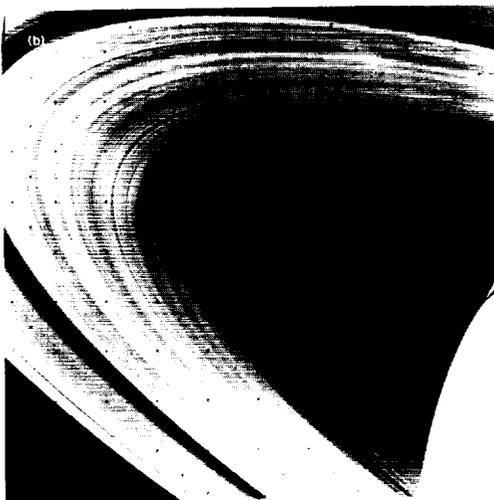
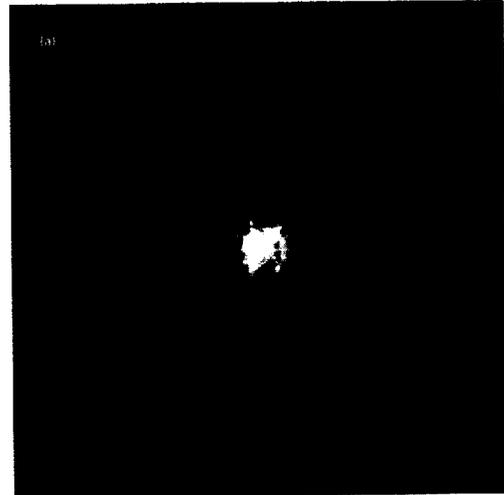
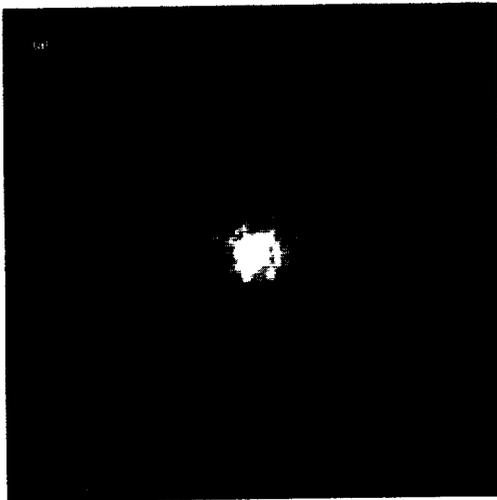


Fig. 1. Original test images for CRAF/Cassini: (a) Hubble image "I2," (b) Saturn image "saturn1," and (c) Saturn Image "saturn2."

Fig. 2. Reconstructed test images for CRAF/Cassini: (a) Hubble image "I2," (b) Saturn image "saturn1," and (c) Saturn Image "saturn2."

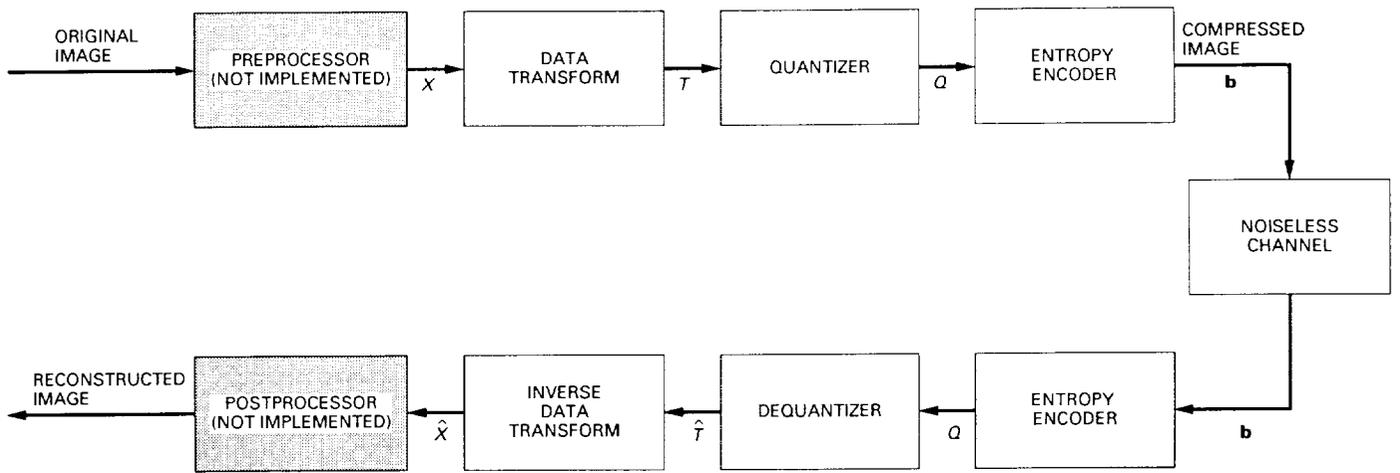


Fig. 3. Block diagram of compression system.

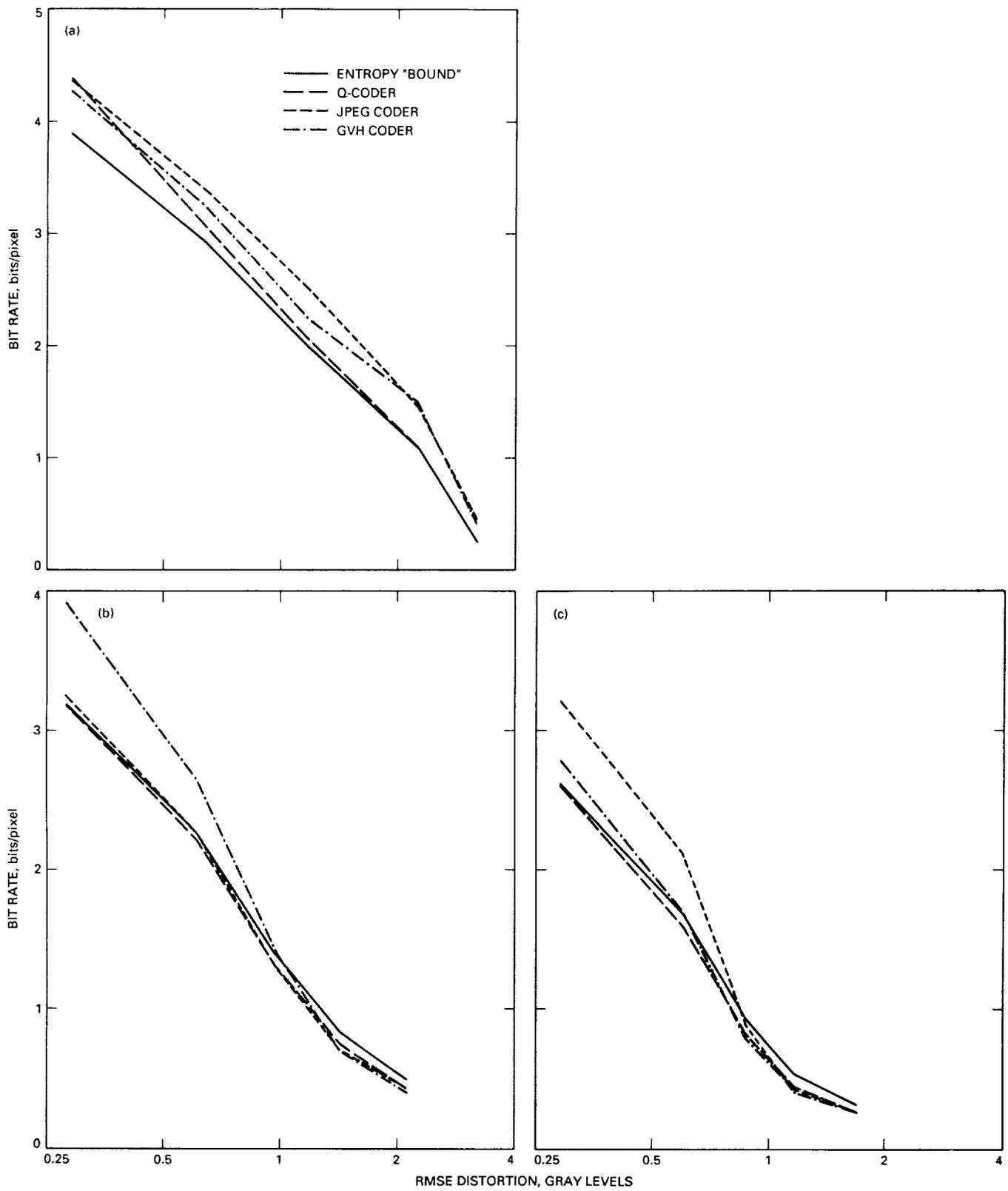


Fig. 4. Bit rate versus RMSE distortion for: (a) Hubble image "l2," (b) Saturn image "saturn1," and (c) Saturn image "saturn2."

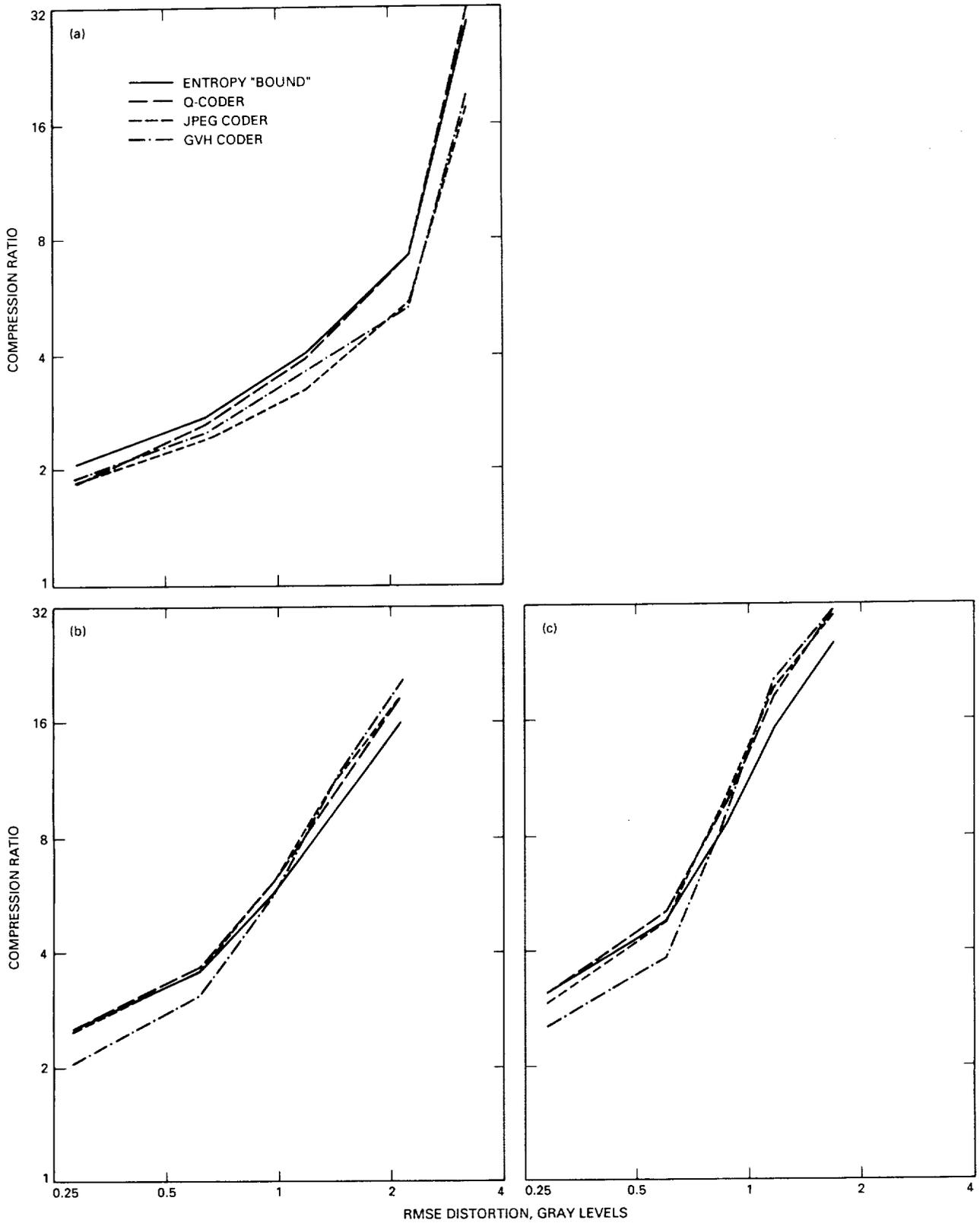


Fig. 5. Compression ratio versus RMSE distortion for: (a) Hubble image "I2," (b) Saturn image "saturn1," and (c) Saturn image "saturn2."

Sg-61  
22.007  
P-5

N91-18317

JJ5-74457

# Emerging Standards for Still Image Compression: A Software Implementation and Simulation Study

F. Pollara and S. Arnold  
Communications Systems Research Section

*This article describes the software implementation of an emerging standard for the lossy compression of continuous-tone still images. This software program can be used to compress planetary images and other two-dimensional instrument data. It provides a high-compression image-coding capability that preserves image fidelity at compression rates competitive or superior to most known techniques. This software implementation confirms the usefulness of such data compression and allows its performance to be compared with other schemes used in deep-space missions and for database storage.*

## I. Introduction

The Joint Photographic Experts Group of the International Standards Organization, together with the International Consultative Committee for Telephone and Telegraph, is in the process of developing an international standard for still image compression with both transmission and storage applications [1]. In its baseline version, the proposed algorithm consists of an  $8 \times 8$  discrete cosine transform (DCT), coefficient quantization, and entropy coding (Huffman or arithmetic). The complete encoder/decoder system is illustrated by the block diagram in Fig. 1. This scheme provides a lossy high-compression image coding capability that preserves image fidelity at compression rates competitive or superior to most known techniques [2]. Its software implementation is discussed in the following section.

## II. Structure of the Software Implementation

Image samples, or pixels, are read from the original image file and sent to a two-dimensional DCT module, which produces 64 coefficients that are independently and uniformly quantized with a different step-size for each coefficient. Then a one-dimensional array is formed by reading the  $8 \times 8$  matrix of quantized coefficients in a zig-zag fashion. The sequence of direct current (dc) components—the first coefficient of each block—is differentially encoded, while the alternating current (ac) components are run-length encoded. Finally, some of the most significant bits of each code are further encoded with a variable-length code; the remaining bits are transmitted essentially intact.

Flow diagrams of the software encoder and decoder structures with Huffman coding are shown in Figs. 2 and 3.

Each of the operations described in the following sections has been implemented as a separate software module to allow for future testing of modified modules.

### A. Discrete Cosine Transform Module

The forward and inverse two-dimensional DCTs used in this software implementation are defined by

$$F(u, v) = C(u)C(v) \frac{1}{4} \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \\ \times \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16}$$

and

$$f(x, y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v)F(u, v) \\ \times \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16}$$

where  $C(k) = 1/\sqrt{2}$  if  $k = 0$ , and  $C(k) = 1$  if  $k \neq 0$ . This definition is efficiently implemented by using row-column decomposition [3]. First, the  $8 \times 1$  DCT of each data column is computed; then the transpose of the resulting matrix is stored as an intermediate result. Finally, the  $8 \times 1$  DCT of each data row is computed to yield the desired two-dimensional DCT. This method has the advantage of considerably reducing the total number of operations required and of limiting to  $(2P - 1)(P - 1)$  (105 for  $P = 8$ ) the number of cosine entries to be stored permanently for a  $P \times P$  transform. A fast very large-scale integration (VLSI) of this method is described in [3]. In this article, the software implementation uses floating-point representation but reduced precision versions have also been considered. Another approach to reducing the computational complexity of this step is to consider other transforms such as the Hadamard transform [4] that can be performed using integer operations with only a slight performance degradation [2].

### B. Quantization Module

The 64 coefficients produced by the forward DCT module are quantized by a uniform or constant step-size quantizer, where the fixed step size may vary from coefficient to coefficient. This is accomplished by a predefined  $8 \times 8$  matrix specifying the step size  $Q(u, v)$  for each coefficient.

The DCT of 8-bit input pixels<sup>1</sup> produces output coefficients with an 8-times-larger range corresponding to a total of 11 bits. The quantization matrix can, of course, be adapted to satisfy different subjective quality measures or different instrument nonlinearities.

### C. Coefficient Modeling Module

Besides using two different Huffman codes for the dc coefficient—the first term in the  $8 \times 8$  matrix of coefficients—and the ac coefficients, these two classes are also differently pre-encoded or modeled.

**1. DC Modeling.** The quantized dc coefficient is differentially pre-encoded by computing its difference with the dc term in the previous block. These differences will then be entropy coded. Their dynamic range has now increased to 12 bits.

Two-dimensional dc prediction, which uses both the previous block's dc term and that of the block above, has also been suggested to take greater advantage of pixel correlation. This feature has not yet been implemented, but it will be included in future revisions of the software.

The prediction residual is then assigned to one of 12 categories  $C_i$  ( $i = 0, \dots, 11$ ) defined by the base-2 logarithm of the residual's absolute value. The resulting 4-bit categories are later Huffman encoded. The remaining information about residual values and sign is encoded by a simple variable-length-integer (VLI) code, in which each codeword is  $C_i$  bits long [1].

**2. AC Modeling.** As a first step, the 63 ac coefficients are reordered into a one-dimensional array by reading the  $8 \times 8$  matrix according to a predefined zig-zag scan path. This reordering ranks the coefficients in approximate order of decreasing magnitude.

The one-dimensional array of ac coefficients is modeled by run-length coding. When a nonzero coefficient is encountered, the number of zeros preceding it and its 4-bit category (one of 11) are concatenated and stored into an 8-bit word for further Huffman coding. The remaining information about the run-length/nonzero ac pair is encoded with the same VLI code discussed above. Since only 4 bits are reserved to represent a run length, only runs up to 15 consecutive zeros can be modeled. Longer runs are artificially broken by transmitting a special code for a run

<sup>1</sup> The software implementation described in this article is designed for easy extension to higher input data precision, up to 12 bits per sample.

length of 16 zeros. Another special code is reserved to signal the end of a block, which also prevents the propagation of eventual channel errors to subsequent blocks.

#### D. Entropy-Coding Module

Entropy coding is the process that actually performs the compaction of the data by reducing statistical redundancy. Either Huffman or arithmetic coding can be used as an entropy-coding method.

**1. Huffman Coding.** The dc category and ac run-length/category pairs are Huffman coded using two different codes. These two codes are not a disjoint partition of a larger prefix code since they contain common codewords. However, they can be decoded by their relative position in the serial stream of codewords, which is known on the receiver side since an end-of-block is always followed by a dc code, which may then be followed by ac codes or by an end-of-block code. This solution allows more efficient encoding with smaller average codeword length at the expense of a slightly more complex synchronization scheme.

The present software implementation uses two default look-up tables for the two Huffman codes. The dc table contains 12 codewords, one for each possible category; the ac table contains 162 codewords, one for each combination of 16 run lengths and 10 categories<sup>2</sup> plus the two special codewords for end-of-block and run-length 16. The maximum length of ac codewords is constrained to 16 bits.

The dc code tables are specified by an array of 12 bytes, which contains a properly ordered list of categories corresponding to a lexicographically ordered list of codewords belonging to a prefix code, and by an array of 16 entries representing the number of codewords of each length. The ac code tables are similarly specified by an array of 162 bytes, which contains a properly ordered list of run-length/category pairs corresponding to a lexicographically ordered list of codewords of a prefix code, and by an array of 16 entries representing the number of codewords of each length. These four arrays completely specify the two codes and can be used to send code information to the decoder or to specify custom tables adapted to the particular

---

<sup>2</sup> The number of categories that can actually occur is 11 (for 8-bit data plus 3-bit expansion due to DCT), but category 0 is unused since only nonzero ac coefficients need to be encoded.

source of interest. This software implementation can also be used to perform a two-pass encoding in which specific codes for the image to be transmitted are created by the encoder during the first pass.

**2. Arithmetic Coding.** As a higher performance alternative to Huffman coding, arithmetic coding has also been included in the standard's specification [1]. Arithmetic coding provides a one-pass scheme that dynamically adapts to image statistics. For this reason, it has a generally superior performance [2] to the nonadaptive Huffman coding chosen for the standard that requires image statistics information before coding. Furthermore, unlike Huffman coding, arithmetic coding does not always require at least one bit per data sample.

### III. Conclusions

This article described a software implementation of a DCT-based lossy compression algorithm suitable for transmitting images from deep space and for storing images in databases. This software is now available for testing and for measuring compression and reproduction-quality performance on various instrument sources of interest. The compression procedure is executed by running the encoder program, which reads the original raster-scanned image—with 8-bit-per-pixel gray-scale resolution—to produce a compressed binary-file image and to compute the compression ratio. Different compression ratios can be obtained by changing the quantization table or the arrays specifying the Huffman codes. The decoder program, in turn, reads the compressed image and produces the reconstructed image in the same format as the original. Very satisfying reproduction quality has been obtained in preliminary tests described in [2].

In the standard considered in this article, practical hardware realization issues have been carefully evaluated to yield designs suitable for VLSI implementation. A commercial VLSI chip set based on this algorithm is already available [5]. Beyond improvements in performance or reductions in complexity that may be possible for specific deep-space instrument applications, the main challenge will be to demonstrate that this algorithm can be realized in a space-qualified version.

## Acknowledgments

The authors acknowledge the useful suggestions of K.-M. Cheung, S. J. Dolinar, and I. Onyszchuk during the development of this software.

## References

- [1] *JPEG Draft Technical Specification (Rev. 5)*, International Standards Organization/International Consultative Committee for Telephone and Telegraph (ISO/CCITT), Washington, D.C., January 15, 1990.
- [2] S. J. Dolinar, K.-M. Cheung, I. Onyszchuk, S. Arnold, and F. Pollara, "Compressed/Reconstructed Test Images for CRAF/Cassini," *TDA Progress Report 42-104* (this issue), vol. October-December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 88-97, February 15, 1991.
- [3] M. Sun, T. Chen, and A. Gottlieb, "VLSI Implementation of a  $16 \times 16$  Discrete Cosine Transform," *IEEE Trans. on Circuits and Systems*, vol. 36, no. 4, pp. 610-617, April 1989.
- [4] W. K. Pratt, *Digital Image Processing*, New York: Wiley & Sons, 1978.
- [5] *Technical Specification Catalog*, LSI Logic Corporation, Milpitas, California, June 1990.

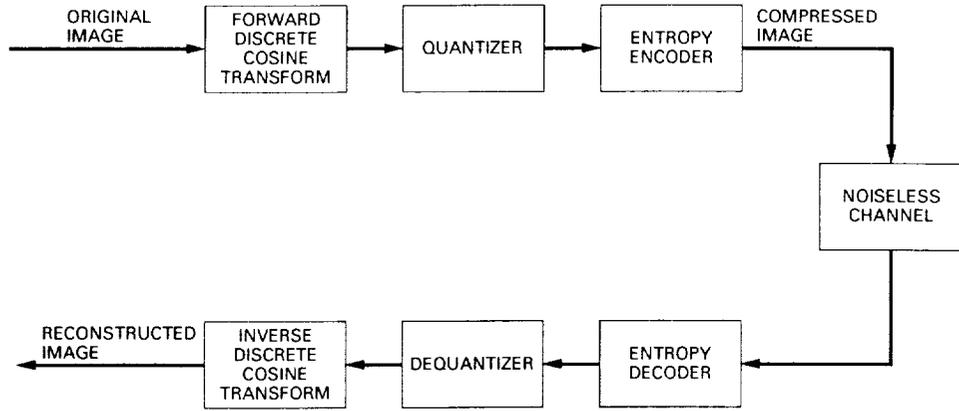


Fig. 1. Block diagram of compression system.

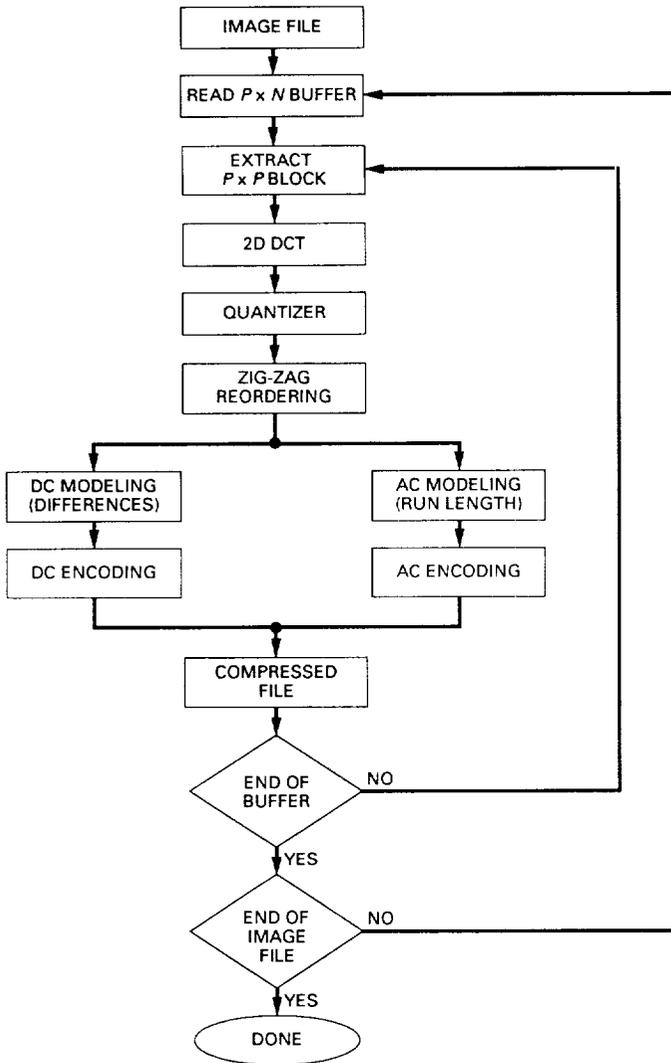


Fig. 2. Baseline compressor for Huffman coding.

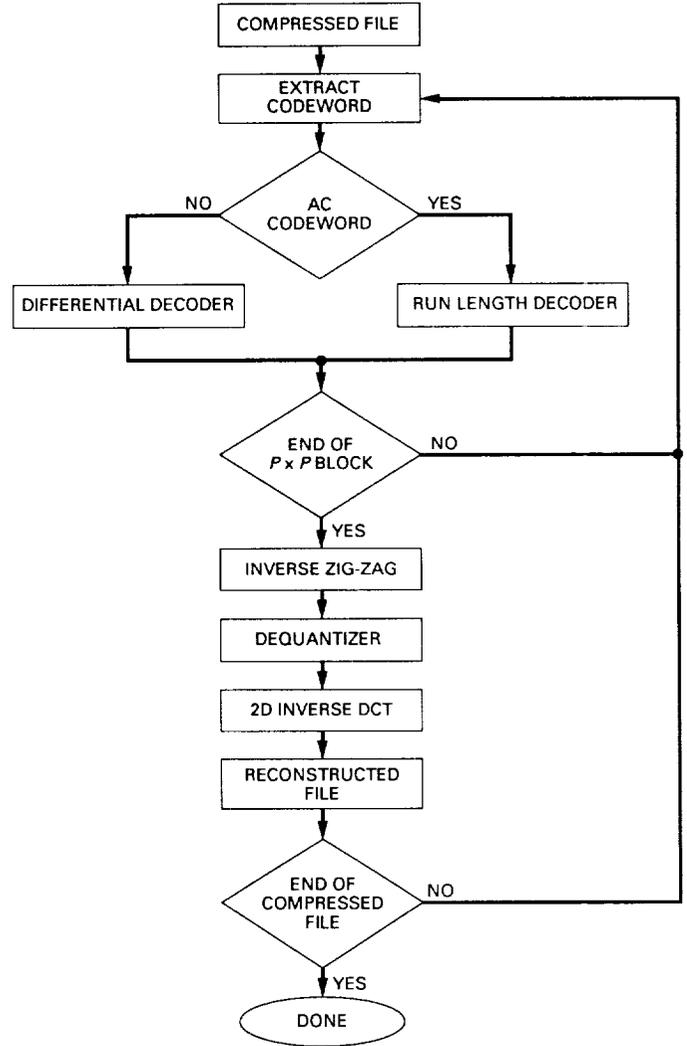


Fig. 3. Baseline decompressor for Huffman coding.

N91-18318

# Algebraic Geometric Codes

M. Shahshahani

Communications Systems Research Section

*This article discusses the performance characteristics of certain algebraic geometric codes. Algebraic geometric codes have good minimum distance properties. On many channels they outperform other comparable block codes; therefore, one would expect them eventually to replace some of the block codes used in communications systems. This article suggests that it is unlikely that they will become useful substitutes for the Reed-Solomon codes used by the DSN in the near future. However, they may be applicable to systems where the signal-to-noise ratio is sufficiently high so that block codes would be more suitable than convolutional or concatenated codes.*

## I. Introduction

In their 1982 paper [1], Tsfasman, Vladut, and Zink showed that, by using algebraic curves, one can construct codes that lie above the Varshamov-Gilbert bound. These codes perform better than other comparable block codes on many channels. This important discovery led to a resurgence of interest in geometric, or Goppa, codes. Most of the research in recent years has focused on developing practical decoding algorithms. For an account of these efforts, refer to [2,3]. This article discusses some of the performance characteristics of certain algebraic geometric codes. For the reader's convenience, some of the basic ideas are developed in Section II as an introduction to the more technical papers in this field.

that is,  $V$  with the origin removed). On  $V^*$  define the equivalence relation

$$[x_0, \dots, x_r] \sim [y_0, \dots, y_r]$$

if  $y_j = \lambda x_j$ , for all  $j$  and some  $\lambda \in \mathbf{F}_q^\times$ . The quotient space  $V^*/\sim$ , where equivalent elements are identified, is the projective space  $\mathbf{P}^r$  over  $\mathbf{F}_q$ . A point  $x \in \mathbf{P}^r$  is represented by  $x = [x_0, \dots, x_r]$  (of course, not uniquely), and  $x_i$ 's are called the homogeneous coordinates of  $x$ . For many geometric problems this space is more convenient to use than the Euclidean space  $\mathbf{F}_q^r$ . There is a way of translating statements about subsets of Euclidean space into those of the projective space, which is described next.

First, embed  $\mathbf{F}_q^r$  into  $V^*$  by

$$i(x_1, \dots, x_r) = [1, x_1, \dots, x_r]$$

Then, to a subset  $S$  of  $\mathbf{F}_q^r$ , assign the image  $\mathbf{S}$  of  $i(S)$  in  $\mathbf{P}^r$ . This means that, in terms of the homogeneous coordinates, i.e., in  $V^*$ ,  $\mathbf{S}$  is represented by the cone

## II. Definition and Basic Properties

Let  $\mathbf{F}_q$  be the field of  $q$  elements ( $q = p^s$ ,  $p$  prime) and  $\mathbf{F}_q^\times$  be the set of nonzero elements of  $\mathbf{F}_q$ . To construct the projective space  $\mathbf{P}^r$  over  $\mathbf{F}_q$ , let  $V = V_{r+1}$  be the vector space of  $(r+1)$ -tuples of elements of  $\mathbf{F}_q$ , and  $V^* = V \setminus \{0\}$

$$S^* = \{[\lambda, \lambda x_1, \dots, \lambda x_r] \mid \lambda \in \mathbf{F}_q^\times, \text{ and } [x_1, \dots, x_r] \in S\}$$

Thus, statements about the subsets of Euclidean space can be translated into statements about those of the projective space. In particular, notice that points in Euclidean space become rays through the origin and that lines in Euclidean space become two-dimensional planes through the origin in  $V^*$ .

To see the value of this translation, consider the special case  $r = 2$ , i.e., the projective plane.  $P^2$  may be regarded as  $\mathbf{F}_q^2$  with a line and a point at infinity added to it. The line is the set  $\{[0, 1, x_2]\}$  and the point is  $[0, 0, 1]$ . Let  $S$  and  $T$  be two parallel lines in the plane  $\mathbf{F}_q^2$ , as defined by the equations

$$S : ax_1 + bx_2 = c \quad \text{and} \quad T : ax_1 + bx_2 = c'$$

Then, the intersection of  $S^*$  and  $T^*$  in  $V^*$  is the line defined by

$$ax_1 + bx_2 = 0 \quad \text{and} \quad x_0 = 0$$

Thus,  $S$  and  $T$  intersect at the point  $[0, 1, -a/b]$  of  $P^2$ . So an important difference between Euclidean space and the projective space is that in the latter space, all lines intersect. Since many problems in geometry can be reduced to problems of intersections, it makes more sense to work in the projective space and avoid the exceptional case of nonintersecting or parallel lines.

Let  $f(x_1, x_2) = \sum a_{ij} x_1^i x_2^j$  be a polynomial in two variables of degree  $e$ . Then, homogenize this polynomial by adding the variable  $x_0$  and considering the homogeneous polynomial

$$F(x_0, x_1, x_2) = \sum a_{ij} x_0^{e-i-j} x_1^i x_2^j$$

For the set  $Z(f) = \{(x_1, x_2) \mid f(x_1, x_2) = 0\}$ , the procedure of going from Euclidean space to the projective space (or to  $V^*$ ) amounts to going from  $Z(f)$  to  $Z(F) = \{[x_0, x_1, x_2] \mid F(x_0, x_1, x_2) = 0\}$ . It is convenient for this application to consider only polynomials  $F$  satisfying a certain technical property (called nonsingularity) that will be described at the end of this section.

Let  $F$  and  $G$  be homogeneous polynomials of degrees  $e$  and  $m$  in three variables. For a subset  $S$ , denote its cardinality by  $|S|$ . Then  $|Z(F) \cap Z(G)|$  is bounded by  $em$ . It is actually equal to  $em$  if the intersections are counted with multiplicities (for example, tangency has multiplicity two, etc.) and allow points to have coordinates in the algebraic closure of  $\mathbf{F}_q$ . These more technical points will not be dis-

cussed here. Note, however, that if  $F$  or  $G$  is a product of linear polynomials, then the assertion that  $|Z(F) \cap Z(G)|$  is bounded by  $em$  follows from the fundamental theorem of algebra.

Now assume that a linear space  $L$  (over  $\mathbf{F}_q$ ) of functions on the subset  $Z(F)$  of  $P^2$  and a subset  $S = \{\xi_1, \dots, \xi_n\}$  of  $Z(F)$  are specified. Consider the mapping

$$\mu : L \rightarrow \mathbf{F}_q^n \quad \text{where} \quad \mu(G) = (G(\xi_1), \dots, G(\xi_n))$$

Then, the image of  $\mu$  is a linear subspace of  $\mathbf{F}_q^n$ , and is, therefore, a code. In order to analyze this code, some control over the linear space  $L$  must be exercised. Here algebraic geometry provides "naturally" defined linear spaces  $L$ , and the parameters of the corresponding code may be evaluated. Note also that certain Reed-Solomon codes may be defined in a similar manner. In fact, if  $L$  is the space of all polynomials of degree less than  $k$  and  $\mathbf{F}_q = \{\xi_1, \dots, \xi_q\}$ , then an extended Reed-Solomon code is the image of the map  $\mu : L \rightarrow \mathbf{F}_q^q$ , where  $\mu(f) = (f(\xi_1), \dots, f(\xi_q))$ . This code has parameters  $(q, k, q - k + 1)$  and is a maximum distance separable code.

Let  $\mathbf{R} = \mathbf{F}_q[x_0, x_1, x_2]$  be the vector space of polynomials in three variables with coefficients in  $\mathbf{F}_q$ , and let  $\mathbf{R}_m$  be the subspace spanned by the homogeneous polynomials of degree  $m$ . In this case (i.e.,  $r = 2$ ),  $Z(F)$  is called a plane curve. It is necessary to construct a linear space of functions on  $Z(F)$  from  $\mathbf{R}_m$ . Notice that for  $G \in \mathbf{R}_m$

$$G(\lambda x_0, \lambda x_1, \lambda x_2) = \lambda^m G(x_0, x_1, x_2)$$

so that  $G$  is not a well-defined function on  $Z(F)$  or  $P^2$ . There are two ways of avoiding this difficulty:

- (1) Define the value of  $G$  at a point  $x$  of  $Z(F)$  or  $P^2$  to be  $G(x_0, x_1, x_2)$  where the representative  $[x_0, x_1, x_2]$  of  $x$  is selected so its first nonzero coordinate is 1.
- (2) Fix a homogeneous polynomial  $H$  of degree  $m$  with the property that  $H(\xi_i) \neq 0$  for all  $\xi_i \in S$ . Then  $G/H$  is a well-defined function on  $Z(F)$ .

With either alternative,  $\mathbf{R}_m$  may be regarded as a linear space  $L_m$  of functions on  $Z(F)$ , and, therefore, the code is denoted by  $C(F, S)$ . One can determine the parameters  $(n, k, d)$  of this code under some additional hypotheses.

Two polynomials  $G$  and  $G' \in \mathbf{R}_m$  determine the same function on  $Z(F)$ , i.e., the same element of  $L_m$  if, and only if, their difference is a multiple of  $F$ . Assuming that the degree  $e$  of  $F$  is less than  $m$ , the dimension of  $L_m$  is expected to be

$$\dim(L_m) = \dim(\mathbf{R}_m) - \dim(\mathbf{R}_{m-e})$$

since multiplication by  $F$  maps  $\mathbf{R}_{m-e}$  into  $\mathbf{R}_m$ . It is easy to see that  $\dim(\mathbf{R}_m) = (1/2)(m+1)(m+2)$ . Substituting and simplifying yields

$$\dim(L_m) = c - g + 1 \quad (1)$$

where the quantities  $c = em$  and  $g = (1/2)(e-1)(e-2)$  are called the degree of  $L_m$  and the genus of the plane curve  $Z(F)$ , respectively. Formula (1) is a very special case of the celebrated Riemann-Roch theorem. The above discussion should take some of the mystery out of this useful formula.

Next, assume that  $n > em$ . To determine the parameters of the code  $C(F, S)$ , suppose that  $\mu(G) = 0$ , then the intersection  $Z(F) \cap Z(G)$  has at least  $n > em$  points. But since  $G$  has degree  $m$  and  $F$  has degree  $e$ ,  $|Z(F) \cap Z(G)| \leq em$ . Therefore,  $G = 0$  and the map  $\mu$  is one to one. This implies that the code  $C(F, S)$  has rate

$$\rho = k/n = \{em - (1/2)(e-1)(e-2) + 1\}/n$$

The minimum distance  $d$  of the code is the minimum number of nonzero entries of  $(G(\xi_1), \dots, G(\xi_n))$  as  $G$  ranges over the nonzero elements of  $L_m$ . As noted,  $|Z(F) \cap Z(G)|$  does not exceed  $em$ , and therefore

$$d \geq n - em$$

Having defined the code  $C(F, S)$ , it is natural to try to understand its dual code  $C(F, S)^*$  with parameters  $(n, k^*, d^*)$ . The computation of the parameters of  $C(F, S)^*$  involves introducing more algebraic geometry, and will not be discussed in detail. The result is:

$$k^* = n - c + g - 1 = n - k \text{ and } d^* \geq c - 2g + 2$$

Here, only a restricted class of algebraic geometric codes was considered. While there are more general constructions, the special case considered will suffice for the problems at hand.

Finally, the nonsingularity property mentioned earlier must be clarified. For a homogeneous polynomial  $F$ ,  $Z(F)$  is nonsingular if, for every  $i$ , the equations

$$\{F_i = 0 \text{ and } \partial F_i / \partial x_j = 0 \text{ for } j \neq i\}$$

do not have a solution. Here,  $F_i$  is the polynomial obtained from  $F$  by setting  $x_i = 1$ . Since for each  $i$ , this is a set of three equations in two unknowns, the nonsingularity condition is generically satisfied. For example, if

the polynomial  $F$  satisfies this condition, then it cannot be written in the form  $F = GH$  with  $G$  and  $H$  homogeneous polynomials of positive degree. In fact, the nonsingularity condition will not be satisfied, since  $F_i = 0$  and  $\partial F_i / \partial x_j = 0$  on  $Z(F) \cap Z(G)$ . On the other hand,  $Z(F)$ , where  $F(x_0, x_1, x_2) = x_0^t + x_1^t + x_2^t$ , is nonsingular for those values of  $t$  and  $p$  which are relatively prime.

### III. Construction and Performance of Certain Codes

It is clear from the expressions for  $k$  and  $d$  that, to construct "good" codes, one should find polynomials with  $|Z(F)|$  as large as possible, so that  $d$  can be large, while the symbol size is fairly small. [Recall that  $|Z(F)|$  means cardinality of  $Z(F)$  as a subset of  $P^2$  and not  $V^*$ .] There is an important inequality (the Weil-Serre bound) relating the size of  $|Z(F)|$  to  $e$ , namely,

$$|Z(F)| \leq q + 1 + g[2\sqrt{q}] \quad (2)$$

where  $[y]$  denotes the largest integer not exceeding  $y$ . This bound is sharp in the sense that there are plane curves for which the equality in Formula (2) is achieved. In order to understand the basic properties of the algebraic geometric code  $C(F, S)$ , compute  $|Z(F)|$ . In the following example,  $|Z(F)|$  is computed for a class of homogeneous polynomials in three variables:

**Example:** Let  $q = p^{ab}$  and  $r = p^a$ , then  $q-1 = (r-1)t$  for some integer  $t$ . Consider the Fermat curve defined by

$$F(x_0, x_1, x_2) = x_0^t + x_1^t + x_2^t \quad (3)$$

For such  $F$

$$|Z(F)| = 3t + (r-2)t^2$$

Consider the mapping  $\chi(\zeta) = \zeta^t$ , which is a homomorphism of  $\mathbf{F}_q^\times$  into itself. Since  $\zeta^{tr} = \zeta^{q-1+t} = \zeta^t$ ,  $\zeta^t$  lies in  $\mathbf{F}_r^\times$ . (Note  $\mathbf{F}_q \supset \mathbf{F}_r$ .) Therefore,  $\chi$  is actually onto  $\mathbf{F}_r^\times$  and its kernel has order  $t$ . First, consider solutions to  $F = 0$  with  $x_0 = 0$ , then one may assume  $x_1 = 1$ . It follows that there are exactly  $3t$  solutions with exactly one coordinate zero. Next, set  $x_0 = 1$ , and seek solutions where all the coordinates are nonzero. For  $-1 \neq \alpha \in \mathbf{F}_r^\times$ , the equation  $\chi(\zeta) = -\alpha - 1$  has  $t$  solutions in  $\mathbf{F}_q^\times$ . Since  $\eta^t = \alpha$  also has  $t$  solutions,  $(r-2)t^2$  solutions were obtained with all the coordinates nonzero. Hence, there are  $3t + (r-2)t^2$  solutions to  $F = 0$  in  $\mathbf{F}_q$ . For  $a = b$  and  $t = r + 1$ ,  $|Z(F)| = r^3 + 1$ . Since  $g = r(r-1)/2$ , and the

right-hand side of Formula (2) is also  $r^3 + 1$ , the inequality of Formula (2) is sharp in the sense described earlier.

This example will suffice for investigating some of the properties of algebraic geometric codes. The general phenomenon is that “good” algebraic geometric codes, comparable in rate and word size to Reed-Solomon codes, have larger minimum distance and smaller symbol size. For example, setting  $a = b$  and  $p = 2$  in the example above, one obtains algebraic geometric codes of rate  $\rho$  with minimum distance  $d$  and word length  $L (= n \times \text{symbol size})$  given by (approximately)

$$d \approx (1 - \rho)2^{3a} \text{ and } L \approx (2a)2^{3a}$$

In fact, note that  $em \gg g$ ,  $n \approx |Z(F)| \approx 2^{3a}$ , and  $d \approx (1 - \rho)n$  to obtain the above estimates. For a Reed-Solomon code of rate  $\rho$

$$d' \approx (1 - \rho)2^N \text{ and } L' \approx n2^N$$

from which the claim follows.

To estimate the parameters of several more specific codes, one considers  $C(F, S)$  where  $F(x_0, x_1, x_2)$  is as in the above example,  $p = 2$ ,  $a = b = 3$ , and  $|Z(F)| = 513$ . Consider the following codes:

- (1)  $C(1)$ :  $n = 504$ ,  $e = 9$ , and  $m = 17$ , so that  $k/n = 1/4$  and  $d = 351$
- (2)  $C(2)$ :  $n = 504$ ,  $e = 9$ , and  $m = 31$ , so that  $k/n = 1/2$  and  $d = 225$
- (3)  $C(3)$ :  $n = 504$ ,  $e = 9$ , and  $m = 52$ , so that  $k/n = 7/8$  and  $d = 36$

The performance of these codes has been studied and compared with that of certain Reed-Solomon codes. Assume that the communication channel is a binary symmetric one, which models a binary-phase shift keyed (BPSK) modulation system over an additive white Gaussian noise channel with a hard limiter (hard decision). It is well-known that a good approximation to the output bit error probability is

$$P \approx (p/s) \sum_{j=u+1, \dots, n} C(n, j) s^j (1-s)^{n-j}$$

where  $n = 504$  for the codes  $C(1)$ ,  $C(2)$ , and  $C(3)$ ,  $C(n, j)$  is the binomial coefficient  $n$  choose  $j$ ,  $u = (1 + d)/2$ , and  $s = 1 - (1 - p)^l$  with  $l$  the symbol size. Recall also that

the information bit signal-to-noise ratio  $E_b/N_0$  is related to the bit error probability  $p$  by the formula

$$p = (1/2) \text{Erfc}[(kE_b/nN_0)^{1/2}]$$

where

$$\text{Erfc}(x) = (2/\pi) \int_x^\infty e^{-t^2} dt$$

The performance of  $C(1)$  and  $C(2)$  is compared with Reed-Solomon codes  $RS(511, 127)$  and  $RS(511, 255)$ , respectively. Note that the comparison is between codes of similar rates, but that the symbols for the Reed-Solomon codes are longer—4599 bits, as compared with 3024 bits for  $C(1)$  and  $C(2)$ . While  $C(1)$  performs somewhat better than  $RS(511, 127)$  (about 1/3 dB better),  $C(2)$  and  $RS(511, 255)$  perform almost identically. The algebraic geometric code  $C(3)$  was compared with the Reed-Solomon code  $RS(255, 223)$ , which is part of the concatenated code used for the Galileo spacecraft. Here again, the performance of the codes is very close, but the Reed-Solomon code has shorter symbol length. Also, note that optimal decoding was assumed in these comparisons. The results are given in Figs. 1–3.

The output bit error was also computed as a function of the channel symbol error probability for  $C(1)$ . The results are given in Fig. 4.

Practical encoding procedures for algebraic geometric codes are not known at this time, and their decoding is more difficult than that of Reed-Solomon codes. While much progress has been made on the decoding of these codes, they cannot always be optimally decoded with algorithms having acceptable complexity (see [2,3]).

## IV. Conclusion

Algebraic geometric codes are “good” block codes. They are the first codes to beat the Varshamov-Gilbert bound, and on many channels outperform other comparable block codes. However, their performance characteristics are such that they are unlikely to be useful for the DSN in the near future. Their most likely application is to systems where the signal-to-noise ratio is sufficiently high so that block codes would be generally more suitable than trellis, convolutional, or concatenated code systems. Like other block codes, algebraic geometric codes can only be hard decoded at this time. When their soft decoding becomes possible, these conclusions will have to be revised.

## Acknowledgments

The author wishes to thank D. Divsalar, S. Dolinar, and F. Pollara for helpful conversations on coding theory. A memo (I.O.M. 331-90.2-042) by Cheung, Dolinar, and Pollara comparing the performance of a rate 1/4 algebraic geometric code with certain Reed-Solomon codes was made available to the author.

## References

- [1] M. A. Tsfasman, S. C. Vladut, and T. Zink, "Modular Curves, Shimura Curves, and Goppa Codes Better Than Varshamov-Gilbert Bound," *Math. Nachr.*, vol. 109, pp. 21–28, 1982.
- [2] J. Justensen, K. J. Larsen, H. E. Jensen, A. Havemose, and T. Hoholdt, "Construction and Decoding of a Class of Algebraic Geometric Codes," *IEEE Transactions on Information Theory*, vol. 35, no. 4, pp. 811–821, July 1989.
- [3] A. N. Skorobogatov and S. G. Vladut, "On the Decoding of Algebraic Geometric Codes," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 1051–1060, September 1990.

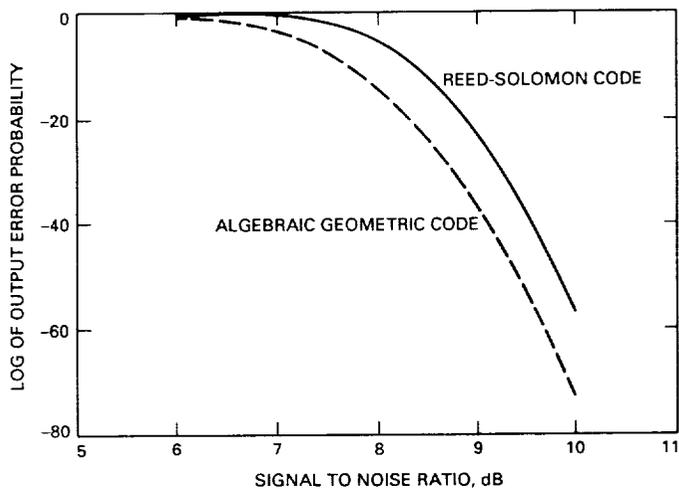


Fig. 1. Comparison of algebraic geometric code  $C(1)$  and Reed-Solomon code (511, 127).

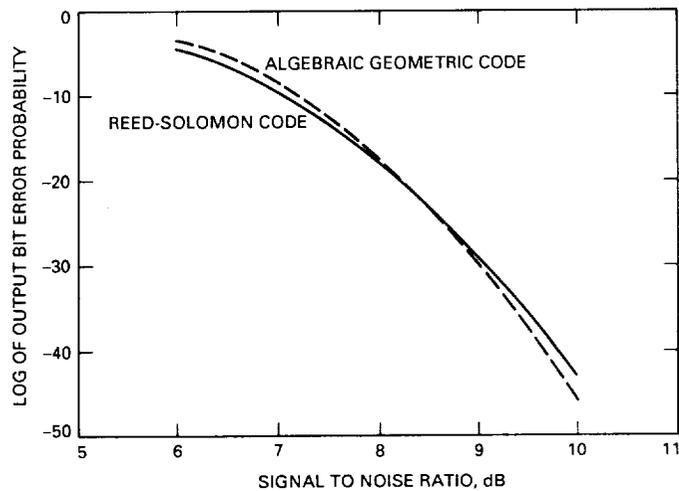


Fig. 3. Comparison of algebraic geometric code  $C(3)$  and Reed-Solomon code (255, 223).

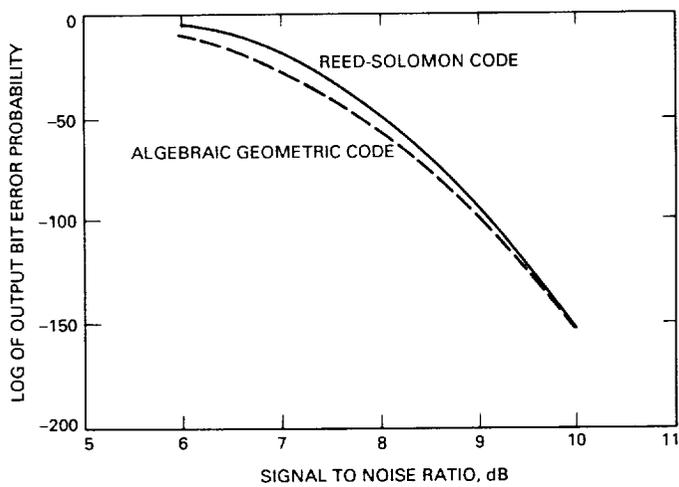


Fig. 2. Comparison of algebraic geometric code  $C(2)$  and Reed-Solomon code (511, 255).

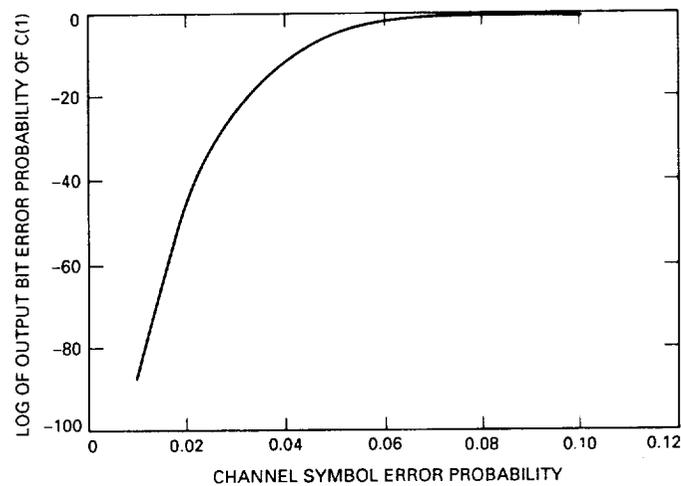


Fig. 4. Output of  $C(1)$ .

511-32  
P-31  
N91-18319  
J J 1090

# Overview of Arraying Techniques in the Deep Space Network

A. Mileant

Telecommunications Systems Section

S. Hinedi

Communications Systems Research Section

*Four different arraying schemes that can be employed by the Deep Space Network are functionally discussed and compared. These include symbol stream combining (SSC), baseband combining (BC), carrier arraying (CA), and full spectrum combining (FSC). In addition, sideband aiding (SA) is also included and compared even though it is not an arraying scheme, since it employs a single antenna. Moreover, combinations of these schemes are discussed, such as carrier arraying with sideband aiding and baseband combining (CA/SA/BC) or carrier arraying with symbol stream combining (CA/SSC). Complexity versus performance is traded off throughout the article and the benefits to the reception of existing spacecraft signals are discussed. Recommendations are made as to the best techniques for particular configurations.*

## I. Introduction

As the signal arriving from a receding deep-space spacecraft becomes weaker and weaker each year, the need arises to devise schemes to compensate for the reduction in signal-to-noise ratio (SNR). With maximum antenna apertures and lower noise temperatures pushed to their limits, the only remaining method to improve the effective SNR is to "combine" the signals from several antennas. This is referred to as arraying and it enables the Deep Space Network (DSN) to extend the missions of spacecraft beyond their planned duration. Another advantage of arraying is its ability to receive higher data rates than can be supported with a single existing antenna. As an example,

symbol stream combining was used to array symbols between the Very Large Array (VLA) radio telescope and Goldstone's antennas during Voyager's encounter at Neptune [1,2]. That technique increased the scientific return from the spacecraft by allowing data transmission at a higher rate. In general, arraying enables a communication link to operate with, effectively, a larger antenna than is physically available.

There are various arraying techniques that have been considered and analyzed in the past. The purpose of this article is to functionally unify and compare the various algorithms and techniques by pointing out their relative ad-

vantages and disadvantages. Antenna arraying can be employed with any signal modulation format: bi-phase shift keyed (BPSK), quadrature phase shift keyed (QPSK), continuous phase modulation (CPM), etc. In this article, the National Aeronautics and Space Administration (NASA) standard deep-space signal format is used to illustrate the different arraying techniques, but the results can be extended to other formats, including suppressed carriers.

It is well known [3] that the received signals from deep-space spacecraft take on the following format:

$$r(t) = \sqrt{2P} \sin[\omega_c t + \Delta d(t) \text{Sin}(\omega_{s,c} t + \theta_{s,c}) + \theta_c] + n(t) \quad (1)$$

where  $n(t)$  is an additive bandlimited white Gaussian noise process,  $P$  is the total received signal power,  $\omega_c$  and  $\theta_c$  are the carrier frequency and phase, respectively,  $\Delta$  is the modulation index,  $d(t)$  is the nonreturn-to-zero or Manchester data, and  $\text{Sin}(\omega_{s,c} t + \theta_{s,c})$  is the square-wave subcarrier with frequency  $\omega_{s,c}$  and phase  $\theta_{s,c}$ . The received signal can be rewritten alternatively as

$$r(t) = \sqrt{2P_C} \sin(\omega_c + \theta_c) + \sqrt{2P_D} d(t) \text{Sin}(\omega_{s,c} t + \theta_{s,c}) \times \cos(\omega_c + \theta_c) + n(t) \quad (2)$$

where  $P_C$  and  $P_D$  are the carrier and data powers and are given by  $P \cos^2 \Delta$  and  $P \sin^2 \Delta$ , respectively. The first component is the residual carrier, typically tracked by a phase-locked loop, and the second component is the suppressed carrier which can be tracked by a Costas loop. The modulation  $d(t)$  is given by

$$d(t) = \sum_{k=-\infty}^{\infty} d_k p(t - kT_s) \quad (3)$$

where  $d_k$  is the  $\pm 1$  binary data and  $T_s$  is the symbol period. The primary function of a receiver is to coherently detect the transmitted symbols as illustrated in Fig. 1. The demodulation process requires carrier, subcarrier, and symbol synchronization. Ideally, the output of the receiver  $x_k$  is given by

$$x_{k,ideal} = \sqrt{P_D} d_k + n_k \quad (4)$$

where  $n_k$  is a Gaussian random variable. In the sections to follow, performance of a particular arraying scheme is measured in terms of its degradation with respect to the ideal gain that can be attained, which assumes no combining or synchronization losses. In the simplest case, using two identical antennas with two separate but identical receivers to demodulate the received signal, ideally the output of each receiver would be similar to Eq. (4) and the noise

samples of each stream would be independent. Therefore, if the symbol streams were properly aligned and added, the symbol SNR of the combined symbol stream would be 3 dB higher than the symbol SNR of the individual streams, resulting in an ideal 3-dB gain.

Typically, however, the carrier, subcarrier, and symbol synchronizations result in signal degradation and the output of the receiver is more realistically modeled by [3]:

$$x_k = \sqrt{P_D} d_k \cos \phi_c \left( 1 - \frac{2}{\pi} |\phi_{s,c}| \right) \left( 1 - \frac{1}{2\pi} |\phi_{s,y}| \right) + n_k \quad (5)$$

where  $\phi_c$ ,  $\phi_{s,c}$ , and  $\phi_{s,y}$  denote carrier, subcarrier, and symbol phase errors, respectively. It is worth noting at this point the difference between symbol SNR degradation and symbol SNR loss. Symbol SNR degradation is defined as the average reduction in SNR at the symbol matched filter output due to imperfect carrier, subcarrier, or symbol synchronization. For example, the SNR degradation due to imperfect carrier reference is given by

$$\overline{C_c^2} = \overline{\cos^2 \phi_c} \quad (6)$$

where the overbar denotes expectation with respect to the carrier phase error  $\phi_c$ . Symbol SNR loss, on the other hand, is defined as the additional symbol SNR needed in the presence of imperfect synchronization to achieve the same symbol error probability as in the presence of perfect synchronization. The latter is typically larger than the SNR degradation, but both degradation and loss are comparable at high loop SNRs [3]. For the purpose of comparing arraying schemes, this article considers symbol SNR degradation, which is easier to compute and is more or less indicative of the error probability performance in the region of interest. In order to simplify this notation, rewrite Eq. (5) as follows:

$$x_k = \sqrt{P_D} C_c C_{s,c} C_{s,y} d_k + n_k \quad (7)$$

where  $C_c$ ,  $C_{s,c}$ , and  $C_{s,y}$  denote the carrier, subcarrier, and symbol "reduction" functions and are given by  $C_c = \cos \phi_c$ ,  $C_{s,c} = 1 - (2/\pi) |\phi_{s,c}|$ , and  $C_{s,y} = 1 - (1/2\pi) |\phi_{s,y}|$ , respectively. Sometimes it is convenient to use  $C_t$  to denote the total degradation given by

$$C_t = C_c C_{s,c} C_{s,y} \quad (8)$$

There are basically four different arraying schemes that can be employed by the DSN. These are symbol stream combining (SSC), baseband combining (BC), carrier arraying (CA), and full spectrum combining (FSC). In addition, sideband aiding (SA) can also be employed, even

though it is not an arraying technique as it employs a single antenna. The next sections will functionally discuss these various schemes and try to clarify their advantages. Furthermore, a combination of these schemes will be discussed, such as carrier arraying with sideband aiding and baseband combining (CA/SA/BC) or carrier arraying with symbol stream combining (CA/SSC), just to name a few. Complexity versus performance will be traded off throughout the article and benefits to the reception of existing spacecraft signals are discussed.

## II. Symbol SNR Degradation Due to Imperfect Synchronization

Before proceeding with arraying, it is crucial to understand and quantify the individual degradations due to the carrier, subcarrier, and symbol synchronizations to assess which of these is the dominant term and how each can be reduced. The results that follow are well known and will be used in subsequent sections to compare arraying schemes.

In the Advanced Receiver II (ARX II) [4], carrier tracking can be performed in two ways. The residual component of the signal can be tracked with a phase-locked loop (PLL) or the suppressed component of the signal can be tracked with a Costas loop. With a PLL, the loop SNR is given by

$$\rho_{c,r} \triangleq \frac{1}{\sigma_{c,r}^2} = \frac{P_C}{N_0 B_c} \quad (9)$$

where  $B_c$  is the carrier loop bandwidth and  $\sigma_{c,r}^2$  is the phase jitter in the loop (the subscript  $c,r$  refers to the carrier residual component). On the other hand, with a Costas loop,

$$\rho_{c,s} \triangleq \frac{1}{\sigma_{c,s}^2} = \frac{P_D S_L}{N_0 B_c} \quad (10)$$

where  $S_L$  is the squaring loss given by

$$S_L = \frac{1}{1 + [1/(2E_s/N_0)]} \quad (11)$$

and  $E_s/N_0 = P_D T_s/N_0$  is the symbol SNR (the subscript  $c,s$  refers to the carrier suppressed component). Note from Eq. (2) that when  $\Delta = 90$  deg, the residual component disappears and the carrier is fully suppressed. On the other hand, when  $\Delta = 0$  deg, the signal reduces to a pure sine wave. Typically both components of the carrier, residual and suppressed, can be tracked simultaneously and the carrier phase estimates combined to provide an improved

estimate. This is referred to as sideband aiding and it results in an improved loop SNR given to a first-order approximation [5] by

$$\rho_c = \rho_{c,r} + \rho_{c,s} \quad (12)$$

Whether sideband aiding is employed or not, the degradation due to imperfect carrier reference is still given by  $\overline{C_c^2}$ . Assuming a carrier phase error density function of the form [6]:

$$p(\phi_c) = \frac{e^{\rho_c \cos \phi_c}}{2\pi I_0(\rho_c)} \quad (13)$$

the carrier degradation,  $\overline{C_c^2}$ , becomes

$$\overline{C_c^2} = \frac{1}{2} \left[ 1 + \frac{I_2(\rho_c)}{I_0(\rho_c)} \right] \quad (14)$$

where  $I_k(x)$  denotes the modified Bessel function of order  $k$ . Note that the Tikhonov density of Eq. (13) is valid only in the case of a first-order PLL. It will, however, be used as an approximation for other synchronization loops, including the subcarrier and the symbol loops, with the carrier loop SNR replaced by the SNR of the respective loop. The degradation  $\overline{C_c^2}$  is shown in Fig. 2 as a function of the ratio of the total received signal power to the one-sided noise spectral level (i.e.,  $P_T/N_0$ ), in the presence and absence of sideband aiding. From the figure, it is clear that sideband aiding can reduce the carrier degradation significantly when the data power is sufficiently "large," i.e., when the modulation index is relatively "high." Furthermore, an "x" has been placed in the figure indicating the point where the carrier loop SNR drops below 8 dB and significant cycle slipping occurs. From Fig. 2, the loop maintains lock at 22.5 dB-Hz with SA, but requires at least 27.5 dB-Hz without sideband aiding, resulting in a 5-dB higher operating threshold.

Another useful quantity is the average amplitude degradation due to imperfect carrier synchronization, given by

$$\overline{C_c} = \frac{I_1(\rho_c)}{I_0(\rho_c)} \quad (15)$$

It will be used in subsequent sections to compute the SNR degradation of various arraying schemes. As for the subcarrier and symbol phase errors, consider two cases, one where the densities are assumed Gaussian, and the other where they are approximated by the Tikhonov density of Eq. (13). Both approximations are expected to agree at high loop SNRs, but not necessarily at low loop SNRs. The exact densities of the phase errors are not known and remain an open problem. For either density,

$$\overline{C_{sc}} = 1 - \frac{2}{\pi} \overline{|\phi_{sc}|} \quad (16)$$

$$\overline{C_{sy}} = 1 - \frac{1}{2\pi} \overline{|\phi_{sy}|} \quad (17)$$

$$\overline{C_{sc}^2} = 1 - \frac{4}{\pi} \overline{|\phi_{sc}|} + \frac{4}{\pi^2} \overline{\phi_{sc}^2} \quad (18)$$

$$\overline{C_{sy}^2} = 1 - \frac{1}{\pi} \overline{|\phi_{sy}|} + \frac{1}{4\pi^2} \overline{\phi_{sy}^2} \quad (19)$$

where

$$\overline{|\phi|} = \sqrt{\frac{2}{\pi}} \sigma \quad (20)$$

$$\overline{\phi^2} = \sigma^2 \quad (21)$$

for the Gaussian densities, and  $\sigma^2$  denotes the variance of the phase error, i.e.,  $\rho = 1/\sigma^2$ . Now, let  $\rho$  denote the loop SNR, then

$$\overline{|\phi|} = \frac{1}{\pi I_0(\rho)} \sum_{k=0}^{\infty} \epsilon_k I_k(\rho) \frac{1}{k^2} [(-1)^k - 1] \quad (22)$$

$$\overline{\phi^2} = \frac{2}{I_0(\rho)} \sum_{k=0}^{\infty} \epsilon_k I_k(\rho) \frac{(-1)^k}{k^2} \quad (23)$$

for the Tikhonov densities ( $\epsilon_0 = 1$  and  $\epsilon_k = 2$  for  $k \neq 0$ ). The subcarrier loop phase jitter,  $\sigma_{sc}^2$ , in a Costas loop is given by ( $W_{sc}$  denotes the subcarrier window) [7]:

$$\sigma_{sc}^2 = \frac{1}{\rho_{sc}} = \left(\frac{\pi}{2}\right)^2 \frac{B_{sc} W_{sc}}{R_s (E_s/N_0)} \left(1 + \frac{E_s}{N_0}\right) \quad (24)$$

Similarly, the symbol loop phase jitter,  $\sigma_{sy}^2$ , assuming a data transition tracking loop, is [6]:

$$\sigma_{sy}^2 = \frac{1}{\rho_{sy}} = 2\pi^2 \frac{B_{sy} W_{sy}}{R_s (E_s/N_0) \operatorname{erf}^2(\sqrt{E_s/N_0})} \quad (25)$$

where  $W_{sy}$  is the symbol window,  $R_s = 1/T_s$  is the symbol rate, and  $\operatorname{erf}(x)$  denotes the error function. The subcarrier and symbol degradations,  $\overline{C_{sc}^2}$  and  $\overline{C_{sy}^2}$ , are depicted in Fig. 3 versus the loop SNR for both approximations.

Typically, the DSN operates with a subhertz loop bandwidth for the subcarrier and symbol synchronization loops,

resulting in negligible degradation from imperfect subcarrier and symbol phase references. In most situations, the carrier degradation is the dominant term and can be as large as 0.8 dB for very weak signals, with an 8-dB loop SNR. Since both approximations yield similar degradations, the performances of the various schemes in the subsequent sections are derived using the Gaussian model for the phase errors.

### III. Arraying Techniques

As mentioned previously, a 3-dB improvement in symbol SNR is expected, assuming an array of two identical antennas with ideal synchronization. In this section, the effective symbol SNR is derived for the various arraying schemes, assuming  $L$  antennas and accounting for imperfect synchronization. Moreover, included here is the effect of sideband aiding and a comparison of performance versus complexity for all schemes.

#### A. Symbol Stream Combining (SSC)

Symbol stream combining is depicted in Fig. 4. Each antenna tracks the carrier and the subcarrier and performs symbol synchronization individually. The symbols at the output of each receiver are then combined, with the appropriate weights, to form the final detected symbols. The advantage of SSC is that the combining loss is negligible [8] and is performed in the data rate bandwidth. Moreover, antennas that are continents apart can transmit their symbols in real or non-real time to a central location where the symbol stream combiner outputs the final symbols. That, however, requires that each antenna be able to lock on the signal individually. The disadvantage of SSC is that  $L$  carrier,  $L$  subcarrier, and  $L$  symbol tracking devices are needed, and each suffers some degradation. For “moderate” to “high” modulation indices, the carrier degradation can be reduced by employing sideband aiding at each antenna. The samples of the signal at the output of the symbol stream combiner are

$$v_k = d_k \sum_{i=1}^L \beta_i \sqrt{P_{Di}} \cos \phi_{ci} \left(1 - \frac{2}{\pi} |\phi_{sci}|\right) \times \left(1 - \frac{1}{2\pi} |\phi_{syi}|\right) + n'_k \quad (26)$$

where  $\beta_i$ 's are weighing factors,  $P_{Di} = P_i \sin^2 \Delta$  is the received data power at antenna  $i$  ( $P_i$  is total received power), and  $\phi_{ci}$ ,  $\phi_{sci}$ , and  $\phi_{syi}$  are the carrier, subcarrier, and symbol phase errors, respectively, at the  $i$ th antenna. There is negligible loss when combining the symbols ( $< 0.05$  dB)

and, assuming that each receiver chain has a one-sided noise power spectral density level  $N_{0i}$ , it is straightforward to show [9] that the variance of  $n'_k$  is given by

$$\sigma_{n'}^2 = \frac{1}{2T_s} \sum_{i=1}^L \beta_i^2 N_{0i} \quad (27)$$

The conditional symbol SNR (assuming that the various phase errors are known) at the output of the combiner is defined as

$$SNR' = \frac{(\bar{v}_k)^2}{\sigma_{n'}^2} \quad (28)$$

where  $\bar{v}_k$  is the conditional mean of  $v_k$ . Using Eq. (27),

$$SNR' = \frac{2P_{D1}T_s}{N_{01}} \frac{\left( \sum_{i=1}^L \beta_i \sqrt{P_{Di}/P_{D1}C_{ti}} \right)^2}{\sum_{i=1}^L \beta_i^2 (N_{0i}/N_{01})} \quad (29)$$

where  $C_{ti}$  is defined in Eq. (8). Letting  $\beta_1 = 1$  and optimizing  $\beta_i$ 's ( $i = 2, \dots, L$ ) in order to maximize  $SNR'$ , one obtains

$$\beta_i = \sqrt{\frac{P_i}{P_1} \frac{N_{01}}{N_{0i}}} \quad (30)$$

Plugging back in Eq. (29), one gets

$$SNR' = \frac{2P_{D1}T_s}{N_{01}} \frac{\left( \sum_{i=1}^L \gamma_i C_{ti} \right)^2}{\sum_{i=1}^L \gamma_i} \quad (31)$$

where  $\gamma_i$  is defined as

$$\gamma_i = \frac{P_i}{P_1} \frac{N_{01}}{N_{0i}} \quad (32)$$

where

$$\Gamma \triangleq \sum_{i=1}^L \gamma_i$$

The  $\gamma_i$  factors for various DSN antennas are given in Appendix A for both S-band (2.2 to 2.3 GHz) and X-band (8.4 to 8.5 GHz). Note that in the absence of any degradation ( $C_{ti} = 1$  for  $i = 1, \dots, L$ ), the conditional SNR simplifies to

$$SNR_{ideal} = \frac{2P_{D1}T_s}{N_{01}} \sum_{i=1}^L \gamma_i = \frac{2P_{D1}T_s}{N_{01}} \Gamma \quad (33)$$

with  $\Gamma$  being the ‘‘ideal gain factor’’ obtained at antenna 1, which will be denoted the master antenna for pure convenience (when  $\gamma_i = \gamma_1$  for all  $i$ ,  $\Gamma = L$ ). For two identical antennas with equal noise temperatures,  $\gamma_1 = \gamma_2 = 1$  and the conditional SNR reduces to  $4P_{D1}T_s/N_{01}$  as expected, i.e., an effective gain of 3 dB. The unconditional SNR at the output of the symbol combiner is obtained by averaging the conditional SNR over the unknown phase errors, which are embedded in the constant  $C_{ti}$ 's, i.e.,

$$SNR = \frac{2P_{D1}T_s}{N_{01}} \left( \frac{\sum_{i=1}^L \gamma_i^2 \overline{C_{ti}^2} + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j \overline{C_{ti} C_{tj}}}{\Gamma} \right) \quad (34)$$

where

$$\overline{C_{ti}^2} = \overline{C_{ci}^2} \overline{C_{sci}^2} \overline{C_{syi}^2} \quad \text{and} \quad \overline{C_{ti}} = \overline{C_{ci}} \overline{C_{sci}} \overline{C_{syi}} \quad (35)$$

Because the noise processes make all the phase errors mutually independent, the computation of the unconditional SNR in Eq. (34) reduces to the computation of the first two moments of the various values of  $C_{ci}$ ,  $C_{sci}$ , and  $C_{syi}$ . Finally, the SNR ‘‘degradation factor’’  $D_{ssc}$  (in decibels) for symbol stream combining is defined as

$$D_{ssc} = 10 \log_{10} \left( \frac{SNR}{SNR_{ideal}} \right) \\ = 10 \log_{10} \left( \frac{\sum_{i=1}^L \gamma_i^2 \overline{C_{ti}^2} + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j \overline{C_{ti} C_{tj}}}{\Gamma^2} \right) \quad (36)$$

Note that  $D_{ssc}$  is a negative number that ideally approaches zero. In general, the larger the  $D_{ssc}$ , the better the symbol stream combining performance. For the case of a single antenna (i.e., no arraying),  $D_{ssc}$  measures the degradation due to imperfect synchronization. Figure 5 depicts  $D_{ssc}$  for the array of two high-efficiency (HEF) and one standard (STD) 34-m antennas as a function of  $P_T/N_0$  of the master antenna, Fig. 5(a), and modulation index, Fig. 5(b). In this case, the ideal expected gain is 2.6 dB, but only a fraction of that is attained, depending on the received  $P_T/N_0$  and  $\Delta$ . The figures also depict the corresponding carrier, subcarrier, and symbol degradations and it is clear that the carrier provides the dominant term.

## B. Baseband Combining (BC)

In baseband combining, each antenna locks on the signal by itself as depicted in Fig. 6. The baseband signals, consisting of data on a subcarrier, are then combined and the symbols demodulated. The combining is performed by the baseband assembly (BBA), which consists of three elements: the real-time combiner (RTC), the subcarrier loop, and the symbol tracking loop. The inputs to the BBA are analog baseband telemetry signals from  $L$  receivers ( $L \leq 8$ ) and the output is a sequence of combined digital symbols given by

$$v_k = d_k \left(1 - \frac{2}{\pi} |\phi_{sc}| \right) \left(1 - \frac{1}{2\pi} |\phi_{sy}| \right) \times \sum_{i=1}^L \beta_i \sqrt{P_{D_i}} \cos \phi_{ci} (1 - 4m |\tau_i|) + n'_k \quad (37)$$

where  $m$  is the ratio of the subcarrier frequency over the symbol rate and  $\tau_i$  is the delay error of the  $i$ th RTC loop ( $\tau_1 \equiv 0$ ) [10]. Since the BBA employs baseband combining (i.e., combines the signals prior to the subcarrier loop), only one subcarrier and one symbol tracking loop are employed and no subscripts are needed for the random variables  $\phi_{sc}$  and  $\phi_{sy}$ . The variance of  $v_k$  due to thermal noise is still given by Eq. (27). Again, as with the SSC scheme, the conditional SNR at the output of the symbol tracking loop is given by

$$SNR' \triangleq \frac{(\bar{v}_k)^2}{\sigma_{n'}^2} = \frac{2P_{D_1}T_s}{N_{01}} \frac{\left( C_{sc}C_{sy} \sum_{i=1}^L \beta_i \sqrt{P_{D_i}/P_{D_1}} C_i \right)^2}{\Gamma} \quad (38)$$

where the SNR degradation function  $C_i$  accounts for the carrier and delay degradations and is defined as

$$C_i = C_{ci}C_{\tau_i} = \cos \phi_{ci} (1 - 4m |\tau_i|) \quad (39)$$

In order to compute the unconditional SNR at the output of the symbol tracking loop, Eq. (38) is averaged over all the phase and delay error processes in the corresponding tracking loops, resulting in

$$SNR = \frac{2P_{D_1}T_s}{N_{01}} \frac{\overline{C_{sc}^2} \overline{C_{sy}^2}}{\Gamma} \left( \frac{\sum_{i=1}^L \gamma_i^2 \overline{C_i^2} + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j \overline{C_i} \overline{C_j}}{\Gamma} \right) \quad (40)$$

The signal reduction function for the RTC, denoted by  $\overline{C_{\tau_i}}$  and given by  $(1 - 4m |\tau_i|)$ , has the following first two moments:

$$\overline{C_{\tau_i}} = \left(1 - 4m \sqrt{\frac{2}{\pi}} \sigma_{\tau_i}\right) \quad (41a)$$

and

$$\overline{C_{\tau_i}^2} = \left(1 - 8m \sqrt{\frac{2}{\pi}} \sigma_{\tau_i} + 16m^2 \sigma_{\tau_i}^2\right) \quad (41b)$$

where  $\sigma_{\tau_i}^2$  denotes the variance of the  $i$ th loop of the real-time combiner and is computed to be [10]:

$$\sigma_{\tau_i}^2 = \frac{B_{\tau_i}}{B_n 32m^2} \times \left\{ \frac{1}{\left[ \operatorname{erf} \left( \sqrt{P_{D_1} \sigma_1^2} \right) \operatorname{erf} \left( \sqrt{P_{D_i} \sigma_i^2} \right) \right]^2} - 1 \right\} \quad (42)$$

In the above equation,  $i = 2, \dots, L$ ,  $B_{\tau_i}$  denotes the bandwidth of the RTC loops,  $B_n$  denotes the noise bandwidth at the RTC input (assumed the same in all channels), and  $\sigma_i^2 = N_{0i} B_n$  (note that  $\overline{C_{\tau_1}} \equiv 1$  and  $\overline{C_{\tau_1}^2} \equiv 1$ ).

The equations for the moments of  $C_{sc}$  and  $C_{sy}$  are those given by Eqs. (16) through (19), with the variances computed using the combined  $P_D/N_0$ . Note that under ideal conditions (i.e., no phase or delay errors in the tracking loops), all  $C$ 's are 1 and the SNR reduces to

$$SNR_{ideal} = \frac{2P_{D_1}T_s}{N_{01}} \Gamma \quad (43)$$

as in the symbol stream combining case, Eq. (33). As expected, both SSC and BC have the same SNR performance under ideal conditions. Once the unconditional SNR is computed for the BC scheme using Eq. (40), the degradation factor is obtained as before, namely,

$$D_{bc} = 10 \log_{10} \left( \frac{SNR}{SNR_{ideal}} \right) = 10 \log_{10} \left( \frac{\overline{C_{sc}^2} \overline{C_{sy}^2}}{\Gamma^2} \frac{\sum_{i=1}^L \gamma_i^2 \overline{C_i^2} + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j \overline{C_i} \overline{C_j}}{\Gamma^2} \right) \quad (44)$$

Figure 7 depicts the degradation due to baseband combining,  $D_{bc}$ , as a function of both  $P_T/N_0$ , Fig. 7(a), and  $\Delta$ , Fig. 7(b), assuming the same array as in the SSC case. The noise bandwidth of the RTC was set to 132 kHz to pass the fifth harmonic of the subcarrier with frequency 32.768 kHz. Note from Fig. 7(a) that the subcarrier and symbol degradations are less than their counterparts in SSC, Fig. 5(a). However, there is an additional combining loss of about 0.2 dB that is not present in the SSC.

### C. Full Spectrum Combining (FSC)

Full spectrum combining is an arraying technique where the signals are combined at intermediate frequency (IF) as depicted in Fig. 8. One receiver chain, consisting of one carrier, one subcarrier, and one symbol synchronization loop, is then used to demodulate the signal. The combining at IF is two dimensional in the sense that both delay and phase adjustment are required to coherently add the signals. Let the received signal at antenna 1 be denoted by  $s_1(t)$ . Then from Eq. (1),

$$s_1(t) = \sqrt{2P_1} \sin[\omega_c t - \theta_1(t)] \quad (45)$$

where  $\theta_1(t) = \theta_M(t) + \theta_D(t) + \theta_{osc}(t)$ ,  $\theta_M(t)$  represents the biphasic modulation,  $\theta_D(t)$  is the Doppler due to spacecraft dynamics, and  $\theta_{osc}(t)$  is the oscillator phase noise.

The received signals at the other antennas are delayed versions of  $s_1(t)$  and are given by

$$s_i(t) = s(t - \tau_i) = \sqrt{2P_i} \sin[\omega_c(t - \tau_i) + \theta_i(t - \tau_i)] \quad (46)$$

for  $i = 2, \dots, L$ , where  $\tau_i$  denotes the delay in signal reception between the first and the  $i$ th antennas ( $\tau_1 \equiv 0$ ), and  $\theta_i(t) = \theta_1(t) + \Delta\theta_i(t)$ . In this case,  $\Delta\theta_i(t)$  accounts for differential Doppler and phase noises, which are typically "very small." Note that at the RF frequency  $\omega_c$ , the signal  $s_i(t)$  can be delayed by  $-\tau_i$  and added coherently, as long as the  $\tau_i$ 's are known. So combining can be achieved at RF with only a delay adjustment. Downconverting the delayed signals to IF ( $\omega_I$  denotes the IF frequency) yields

$$y_i(t) = \sqrt{2P_i} \sin[\omega_I t - \omega_c \tau_i + \theta_i(t - \tau_i)] \quad (47)$$

and delaying each signal  $y_i(t)$  by  $-\tau_i$ , gives

$$\begin{aligned} z_i(t) &= y_i(t + \tau_i) \\ &= \sqrt{2P_i} \sin[\omega_I t - \omega_c \tau_i + \omega_I \tau_i + \theta_1(t) + \Delta\theta_i(t)] \end{aligned} \quad (48)$$

for  $i \neq 1$ . The signal's  $z_i(t)$ 's cannot be added coherently because the phases are not aligned, due to the factor  $(\omega_I - \omega_c)\tau_i$ , even though the data are aligned ( $\theta_D(t)$  is part of  $\theta_1(t)$ ). Therefore, an additional phase adjustment is necessary to add the signals coherently. This example illustrates that both delay and phase adjustments are required to add the signals coherently at IF, but that only a delay compensation is sufficient at RF. For the purpose of this article, a delay by  $-\tau$  (actually, an advancement) is used for mathematical convenience. In reality, the "furthest" antenna can be used as a reference and signals from all other antennas can be delayed accordingly.

Now consider an antenna interferometric pair as illustrated in Fig. 9. The signal at antenna  $i$  arrives  $\tau_i$  sec later than the signal at antenna 1, which will be used as a reference for mathematical convenience. After low noise amplification, the signals are downconverted to IF, where the  $i$ th signal is delayed by  $-\tau_i$  sec. The latter delay consists of two components, a fixed component and a time-varying component. The fixed component compensates for unequal waveguide lengths between the two antennas and the correlator. It is a known quantity that can be determined by measurement. The time-varying component compensates for unequal propagation length for the two received signals. This component is typically precomputed from the trajectory of the spacecraft and the physical location of the two antennas.

The relative phase difference between the signals is estimated by performing a correlation on the resulting signals, which for all practical purposes have been aligned in time. At the input to the correlator, the two signals from the first and the  $i$ th antennas are passed through filters with bandwidth  $B$  Hz and subsequently sampled at the Nyquist rate of  $2B$  samples per sec. Mathematically, the sampled signals are given by

$$z_1(t_k) = \sqrt{2P_1} \sin[\omega_I t_k + \theta_1(t_k)] + n_1(t_k)$$

and

$$z_i(t_k) = \sqrt{2P_i} \sin[\omega_I t_k + (\omega_I - \omega_c)\tau_i + \theta_i(t_k)] + n_i(t_k) \quad (49)$$

where  $i \neq 1$  and  $n_1(t_k)$  and  $n_i(t_k)$  are independent Gaussian random variables with variances  $\sigma_1^2 = N_{01}B$  and  $\sigma_i^2 = N_{0i}B$ . It will be shown later that the parameter  $B$  is essential in determining the averaging period and, thus, the combining loss. Correlating the signals (i.e., multiplying and lowpass filtering) yields

$$I_{i1}(t_k) = \sqrt{P_1 P_i} \cos[\phi_{i1}(t_k)] + n_{I,i1}(t_k) \quad (50)$$

where  $\phi_{i1} = (\omega_I - \omega_c)\tau_i + \Delta\theta_i(t_k)$  denotes the total phase difference between the signals and  $n_{I,i1}$ , the effective noise, is given by

$$\begin{aligned} n_{I,i1} &= \sqrt{2P_i} \sin[\omega_I t_k + \phi_{i1} + \theta_i(t_k)] n_1(t_k) \\ &+ \sqrt{2P_1} \sin[\omega_I t_k + \theta_1(t_k)] n_i(t_k) + n_1(t_k) n_i(t_k) \end{aligned} \quad (51)$$

with effective variance

$$\begin{aligned} \sigma_{ei}^2 &= \sigma_1^2 P_i + \sigma_i^2 P_1 + \sigma_1^2 \sigma_i^2 \\ &= B(N_{01} P_i + N_{0i} P_1) + N_{01} N_{0i} B^2 \end{aligned} \quad (52)$$

The correlation is performed in a complex manner (i.e., four real correlations) resulting in an additional signal  $Q_{i1}(t_k)$  given by

$$Q_{i1}(t_k) = \sqrt{P_1 P_i} \sin[\phi_{i1}(t_k)] + n_{Q,i1}(t_k) \quad (53)$$

The noise samples  $n_{I,i1}(t_k)$  and  $n_{Q,i1}(t_k)$  are uncorrelated with identical variances as given by Eq. (52). The correlator output can be represented more conveniently in a complex form as

$$\mathbf{z}_{i1}(t_k) = I_{i1}(t_k) + jQ_{i1}(t_k) \quad (54)$$

Following the correlation, an averaging operation over  $T$  sec is performed to reduce the noise effect. In that period,  $N = 2BT$  independent samples are used to reduce the variance by a factor of  $N$ . The SNR at the output of the accumulator,  $SNR_{i1}$ , is thus given by

$$SNR_{i1} = \frac{NP_1 P_i}{\sigma_{ei}^2} = \frac{P_1}{N_{01}} \frac{2T}{[1 + 1/\gamma_i + (BN_{0i}/P_i)]} \quad (55)$$

where  $\gamma_i$  is as given in Eq. (32). Note that in radiometric applications [10], the SNR is defined as the ratio of the standard deviation of the signal to that of the noise, and is the square root of the SNR defined in the above equation. Assuming that the correlation bandwidth  $B$  is "very large" (in the MHz range), the signal multiplied by the noise term ( $P_1 \sigma_i^2 + P_i \sigma_1^2$ ) can be ignored and the effective noise variance is dominated by the noise multiplied by the noise term ( $\sigma_1^2 \sigma_i^2$ ), i.e.,

$$\sigma_{ei}^2 \simeq \sigma_1^2 \sigma_i^2 \quad (56)$$

and, hence, the SNR can be well approximated by

$$SNR_{i1} \simeq \frac{P_1}{N_{01}} \frac{P_i}{N_{0i}} \frac{2T}{B} \quad (57)$$

An estimate of  $\phi_{i1}$ ,  $\hat{\phi}_{i1}$  can be obtained by computing the inverse tangent of the real and imaginary parts of  $\mathbf{z}_{i1}$ , i.e.,

$$\hat{\phi}_{i1}(t_k) = \tan^{-1} \left[ \frac{Q_{i1}(t_k)}{I_{i1}(t_k)} \right] \quad (58)$$

The probability density function of such an estimate is given in [11] as

$$p(\hat{\phi}_{i1}) = \frac{1}{2\pi} e^{-SNR_{i1}/2} \left\{ 1 + Ge^{G^2} \sqrt{\pi} [1 + \text{erf}(G)] \right\} \quad (59)$$

where  $\text{erf}(x)$  is the error function and

$$G = \sqrt{\frac{SNR_{i1}}{2}} \cos(\hat{\phi}_{i1} - \phi_{i1}) \quad (60)$$

The density in Eq. (59) is plotted in Fig. 10 and its derivation assumes that the noises  $n_{I,i1}$  and  $n_{Q,i1}$  are Gaussian. Even though they are not Gaussian in the strict sense, a Gaussian approximation is still justified by invoking the central limit theorem due to the averaging over  $N$  samples. Figure 10 clearly indicates that a reasonably "good" phase estimate can be obtained for  $SNR_{i1}$  as low as 6 dB. At a moderately high  $SNR_{i1}$ , the distribution can be approximated by a Gaussian distribution with variance

$$\sigma_{\Delta\phi_{i1}}^2 = \frac{1}{SNR_{i1}} \quad (61)$$

An improvement in phase error estimation can be obtained by performing global phasing between  $L$  antennas, which involves  $L(L-1)/2$  complex correlations as the signal from each antenna is correlated with the signal from every other antenna. In the simplest form, the signal from antenna 1 is correlated with all other signals and the phase errors are estimated. Global phasing reduces the residual phase error variance of Eq. (61) by a factor of  $L-2$  by employing least squares calculations [11] ([2] states that the actual reduction is approximately  $0.5(L-2)$ , which means that global phasing gives an advantage over the conventional scheme only for  $L > 4$ ). In addition to global phasing, closed-loop techniques can be utilized to reduce the phase error as illustrated in Appendix B.

**1. Combining Loss of FSC.** In order to compute the combining loss of FSC, consider the IF signals after phase compensation, i.e.,

$$\mathbf{z}_i(t_k) = \sqrt{P_i} e^{j[\omega_i t_k + \theta_i(t_k) + \Delta\phi_{i1}(t_k)]} + \mathbf{n}_i(t_k) e^{j[\omega_i t_k + \theta_i(t_k) + \Delta\phi_{i1}(t_k)]} \quad (62)$$

where  $\Delta\phi_{i1} = \hat{\phi}_{i1} - \phi_{i1}$  refers to the residual phase error between antenna 1 and the  $i$ th signal and  $\mathbf{n}_i(t_k)$  is the complex envelope of the thermal noise with two-sided noise spectral density  $N_{0i}$ . The signal combiner performs the weighted sum of  $\mathbf{z}_i(t_k)$ , namely,

$$\mathbf{z}(t_k) = \sum_{i=1}^L \beta_i \mathbf{z}_i(t_k) = \sum_{i=1}^L \beta_i \left\{ \sqrt{P_i} e^{j[\omega_i t_k + \theta_i(t_k) + \Delta\phi_{i1}(t_k)]} + \mathbf{n}_i(t_k) e^{j[\omega_i t_k + \theta_i(t_k) + \Delta\phi_{i1}(t_k)]} \right\} \quad (63)$$

Note that the variance of the combined complex signal  $\mathbf{z}(t_k)$  is

$$\sigma_z^2 = 2B \sum_{i=1}^L \beta_i^2 N_{0i} \quad (64)$$

The total signal power at the output of the combiner conditioned on residual phases,  $\Delta\phi_{i1}(t_k)$ , is thus given by

$$\begin{aligned} P'_z &= \left[ \overline{\mathbf{z}(t_k)} \right]^2 = \sum_{i=1}^L \sum_{j=1}^L \beta_i \beta_j \sqrt{P_i P_j} C_{IF_i} C_{IF_j}^* \\ &= \sum_{i=1}^L \beta_i^2 P_i C_{IF_i} C_{IF_i}^* + \sum_{i=1}^L \sum_{\substack{j=1 \\ j \neq i}}^L \beta_i \beta_j \sqrt{P_i P_j} C_{IF_i} C_{IF_j}^* \end{aligned} \quad (65)$$

where

$$C_{IF_i} \triangleq e^{j\Delta\phi_{i1}(t_k)} \quad (66)$$

is the complex signal reduction function due to phase misalignment between the  $i$ th and first signals. Assuming that the ensemble average of the phase difference between any two antennas is independent of which antenna pair is chosen and that the residual phase of each antenna pair is Gaussian distributed with variance  $\sigma_{\Delta\phi_{i1}}^2$ , then it can be shown that

$$\begin{aligned} \overline{C_{IF_i} C_{IF_j}^*} &\triangleq C_{ij} = \mathcal{E}\{e^{j[\Delta\phi_{i1}(t_k) - \Delta\phi_{j1}(t_k)]}\} \\ &= \begin{cases} e^{-\frac{1}{2}[\sigma_{\Delta\phi_{i1}}^2 + \sigma_{\Delta\phi_{j1}}^2]} & i \neq j, \sigma_{\Delta\phi_{i1}}^2 \equiv 0 \\ 1 & i = j \end{cases} \end{aligned} \quad (67)$$

Performing the above averaging operation over  $P'_z$ , the total signal power is obtained, namely,

$$P_z = P_1 \left( \sum_{i=1}^L \gamma_i^2 + \sum_{i=1}^L \sum_{\substack{j=1 \\ j \neq i}}^L \gamma_i \gamma_j C_{ij} \right) \quad (68)$$

Note that in an ideal scenario (i.e., no degradation), the signal reduction functions approach 1 ( $C_{ij} = 1 \forall i, j$ ) and Eq. (68) reduces to  $P_z = P_1 L^2$ . Simultaneously, the noise variance of Eq. (64) becomes proportional to  $L$  and, hence, the SNR increases linearly with  $L$  as expected.

**2. Telemetry Performance of FSC.** With full spectrum combining, only one receiver, one subcarrier, and symbol tracking loops are required. The samples of the signal at the output of the integrate-and-dump filter can be expressed as

$$v_k = d_k \sqrt{P_D} \cos \phi_c \left( 1 - \frac{2}{\pi} |\phi_{sc}| \right) \left( 1 - \frac{1}{2\pi} |\phi_{sy}| \right) + n'_k \quad (69)$$

where  $P_D$  is the combined data power given by  $P_z \sin^2 \Delta$  and  $n'_k$  is Gaussian with variance given by Eq. (27).

Repeating now the same steps as with either BC or SSC, it can be shown that the symbol SNR in terms of  $P_{D1} = P_1 \sin^2 \Delta$  is given by

$$SNR = \frac{2P_{D1}T_s}{N_{01}} \frac{1}{C_c^2} \frac{1}{C_{sc}^2} \frac{1}{C_{sy}^2} \left( \frac{\sum_{i=1}^L \gamma_i^2 + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j C_{ij}}{\Gamma} \right) \quad (70)$$

where the loop losses are computed using the combined power, being carrier or data. Note that in the ideal case, Eq. (70) reduces to  $SNR_{ideal}$  of Eq. (33) as expected. The degradation factor for the full spectrum combining scheme,  $D_{jsc}$ , is given, as before, by the ratio (in decibels) of the combined symbol SNR to the ideal symbol SNR, i.e.,

$$\begin{aligned} D_{jsc} &= 10 \log_{10} \left( \frac{SNR}{SNR_{ideal}} \right) \\ &= 10 \log_{10} \left( \frac{\sum_{i=1}^L \gamma_i^2 + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j C_{ij}}{C_c^2 C_{sc}^2 C_{sy}^2 \Gamma^2} \right) \end{aligned} \quad (71)$$

As an example, let  $P_i = P_1$ ,  $N_{0i} = N_{01}$ , and  $\beta_i = 1$  for all antennas, then the signal and noise powers of the real process at the output of the combiner become, respectively,

$$P_z = P_1 \left[ L + 2(L-1)e^{-(\sigma_{\Delta\phi}^2)/2} + (L-2)(L-1)e^{-\sigma_{\Delta\phi}^2} \right]$$

$$\sigma_z^2 = BLN_{01} \quad (72)$$

and the SNR at the combiner's output becomes

$$SNR_z = \frac{P_z}{\sigma_z^2}$$

$$= \frac{P_1 \left[ L + 2(L-1)e^{-(\sigma_{\Delta\phi}^2)/2} + (L-2)(L-1)e^{-\sigma_{\Delta\phi}^2} \right]}{LN_{01}B} \quad (73)$$

With perfect alignment (i.e.,  $\sigma_{\Delta\phi}^2 \rightarrow 0$ ),  $SNR_z$  reduces to

$$SNR_{z,ideal} = \frac{P_1 L}{N_{01} B} \quad (74)$$

as expected and, hence, the combining degradation for the FSC scheme is given by

$D_{fsc} =$

$$10 \log_{10} \left[ \frac{L + 2(L-1)e^{-(\sigma_{\Delta\phi}^2)/2} + (L-2)(L-1)e^{-\sigma_{\Delta\phi}^2}}{L^2} \right] \quad (75)$$

Figure 11 depicts the degradation of FSC,  $D_{fsc}$ , for the same three-element array, as a function of  $P_T/N_0$  of the master antenna, Fig. 11(a), and modulation index  $\Delta$ , Fig. 11(b). It is clear from Fig. 11(a) that the FSC carrier degradation is significantly reduced over those of SSC and BC. Furthermore, the subcarrier and symbol degradations are identical to those of BC and both are much smaller than the SSC degradations, as expected. The primary advantage of FSC is that the carrier loop SNR for this particular array does not decrease below 8 dB and, hence, cycle slipping is not a major issue even at  $P_T/N_0 = 20$  dB-Hz, with a 3-Hz carrier loop bandwidth.

The major drawback of FSC in this example is the long integration time required to maintain a relatively small combining loss. With  $B = 2 \times 135$  kHz (IF bandwidth) and  $T/B = 0.0008$  sec<sup>2</sup>, the integration time is 216 sec,

which is too long for the phase of interest to remain a constant in a practical scenario. At  $P_T/N_0 = 20$  dB-Hz, the SNR at the correlator output, Eq. (57), is roughly 12 dB. In order to reduce the integration length, the correlator bandwidth can be adjusted to pass only the first harmonic of the subcarrier (i.e.,  $B = 2 \times 33$  kHz), resulting in a shorter integration time of 53 sec (still relatively long) and a slight degradation in correlator SNR. Note that even though the correlation is performed with only the first subcarrier harmonic, the combining should be accomplished with the full data spectrum.

#### D. Carrier Arraying (CA)

In carrier arraying, several carrier tracking loops are coupled in order to enhance the received carrier signal-to-noise ratio and, hence, decrease the telemetry ("radio") loss due to imperfect carrier synchronization. The coupling can be performed using phase-locked loops (PLLs) for residual carriers or Costas loops for suppressed BPSK carriers. Only the PLL case is considered in this article to illustrate the idea of carrier arraying. A general block diagram is shown in Fig. 12 where two carrier loops share information to jointly improve their performance as opposed to tracking individually. Carrier arraying by itself does not combine the data and, thus, needs to operate with baseband combining or symbol stream combining to array the telemetry. This is shown in Fig. 13 where baseband combining is employed to array the data spectra.

There are basically two scenarios where one would employ carrier arraying. In the first scenario, a "large" antenna locks on the signal by itself and then helps a "smaller" antenna track. In this case, the signal might experience dynamics requiring a large loop bandwidth and, hence, the signal has to be strong enough to enable the carrier loop to operate with the large bandwidth. A large antenna with a strong signal is first used to track the signal and then the dynamics of the signal are estimated and removed from the weaker signal to enable the other carrier loop to operate with a smaller bandwidth and, hence, a higher loop SNR. In the second scenario, the signal is too weak to be tracked by any single antenna but can be tracked jointly by two or more antennas. The combining methods used in the latter case are similar to those employed in FSC when aligning the phases of pure tones (hence, requiring a smaller correlator bandwidth). In either scenario, carrier arraying can be implemented in one of two ways: at baseband or at an intermediate frequency (IF). Both implementations are discussed in the next sections.

**1. Baseband Carrier Arraying Scheme.** Baseband carrier arraying is illustrated in Fig. 13 where the error

signals at the output of the phase detectors are combined at baseband. This scheme is analyzed in [12] where it is shown that the variance of the phase jitter process in the master PLL is given by

$$\begin{aligned} \sigma_{c1}^2 &= \frac{1}{2\pi j} \oint \left[ \frac{H_1(z)}{1 + \sum_{i=2}^L \gamma_i H_1(z)[1 - H_i(z)]} \right]^2 \frac{dz}{z} \frac{N_{01}}{2T_{c1}P_{c1}} \\ &+ \sum_{i=2}^L \gamma_i^2 \frac{1}{2\pi j} \oint \left[ \frac{H_1(z)[1 - H_i(z)]}{1 + \sum_{i=2}^L \gamma_i H_1(z)[1 - H_i(z)]} \right]^2 \frac{dz}{z} \frac{N_{0i}}{2T_{ci}P_{ci}} \end{aligned} \quad (76)$$

where  $H_i(z)$  is the closed-loop transfer function of the  $i$ th loop,  $T_{ci}$  is the loop update time, and  $\gamma_i$  is as defined in Eq. (32). The above integral is difficult to evaluate in general. However, when  $B_{ci} \ll B_{c1}$  for  $i = 2, \dots, L$ , which is the preferred mode of operation, the above integral can be approximated by

$$\sigma_{c1}^2 \approx \frac{B_{c1} \sum_{i=1}^L \beta_i^2 N_{0i}}{P_{c1} \Gamma^2} \quad (77)$$

which assumes ideal performance. In this case, the master loop SNR becomes

$$\rho_{c1} = \frac{P_{c1}}{B_{c1} N_{01}} \Gamma \quad (78)$$

assuming identical noise spectral densities. The actual variance will typically be larger and requires the evaluation of Eq. (76), which depends on the actual loop filters implemented.

**2. IF Carrier Arraying Scheme.** One form of IF carrier arraying is depicted in Fig. 14 and is conceptually the same as full spectrum combining. In this case, the total power,  $P_i$ , is substituted for by the carrier power,  $P_{ci}$ . So, all equations and results derived in Section III.C regarding the combining loss can be automatically applied to the IF carrier arraying scheme. Phase estimation in this case can be performed by downconverting the received IFs to baseband using a precomputed model of the received Doppler and Doppler rate. The correlation can be computed at baseband using “very small” bandwidths  $B$  and, hence, requiring “short” integration times  $T$ . From Eq. (52), the variance of the  $i$ th carrier correlator is

$$\begin{aligned} \sigma_{ci}^2 &= B(N_{01}P_{ci} + N_{0i}P_{c1}) + N_{01}N_{0i}B^2 \\ &\approx N_{01}N_{0i}B^2 \end{aligned} \quad (79)$$

while the correlator’s SNR is

$$SNR_{c, i1} \approx \frac{P_{c1}}{N_{01}} \frac{P_{ci}}{N_{0i}} \frac{2T}{B} \quad (80)$$

Note that for IF carrier arraying, the bandwidth  $B$  is much narrower than for full spectrum combining since the data spectrum is not employed.

The signal combiner performs the weighted sum of carrier signals  $c_i(t)$ , giving the complex combined carrier signal

$$\begin{aligned} c(t) &= \sum_{i=1}^L \beta_i \left[ \sqrt{P_{ci}} e^{j[\omega_{ci}t + \theta_i(t) + \Delta\phi_{c,i}(t)]} \right. \\ &\quad \left. + n_i(t) e^{j[\omega_{ci}t + \theta_i(t) + \Delta\phi_{c,i}(t)]} \right] \end{aligned} \quad (81)$$

Following Eqs. (65) through (68), the average carrier power and the variance of the combined complex carrier signal  $c(t)$  are, respectively,

$$\begin{aligned} P_c &= \sum_{i=1}^L \beta_i^2 P_{ci} + \sum_{i=1}^L \sum_{\substack{k=1 \\ i \neq k}}^L \beta_i \beta_k \sqrt{P_{ci} P_{ck}} C_{c, ik} \\ &= P_{c1} \left( \sum_{i=1}^L \gamma_i^2 + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j C_{c, ij} \right) \end{aligned} \quad (82)$$

and

$$\sigma_c^2 = 2B \sum_{i=1}^L \beta_i^2 N_{0i} \quad (83)$$

where

$$\begin{aligned} C_{c, ij} &= \varepsilon \{ e^{j[\Delta\phi_{c,i1}(t_k) - \Delta\phi_{c,j1}(t_k)]} \} \\ &= \begin{cases} e^{-\frac{1}{2}[\sigma_{\Delta\phi_{c,i1}}^2 + \sigma_{\Delta\phi_{c,j1}}^2]} & i \neq j, \sigma_{\Delta\phi_{c,i1}}^2 \equiv 0 \\ 1 & i = j \end{cases} \end{aligned} \quad (84)$$

and

$$\sigma_{\Delta\phi_{c,ij}}^2 = \frac{1}{SNR_{cij}} \quad (85)$$

To illustrate the results with a simple example, let  $P_{ci} = P_{c1}$ ,  $N_{0i} = N_{01}$ , and  $\beta_i = 1$  for all antennas. Then the signal and noise powers of the real process at the output of the carrier combiner become, respectively,

$$P_c = P_{c1} \left[ L + 2(L-1)e^{-(\sigma_{\Delta\phi, c}^2)/2} + (L-2)(L-1)e^{-\sigma_{\Delta\phi, c}^2} \right]$$

and

$$\sigma_{\Delta\phi, c}^2 = BLN_{01} \quad (86)$$

resulting in a correlator SNR

$$SNR_c = \frac{P_c}{\sigma_{\Delta\phi, c}^2} = \frac{P_{c1} \left[ L + 2(L-1)e^{-(\sigma_{\Delta\phi, c}^2)/2} + (L-2)(L-1)e^{-\sigma_{\Delta\phi, c}^2} \right]}{LN_{01}B} \quad (87)$$

In an ideal scenario,  $\sigma_{\Delta\phi, c}^2 \rightarrow 0$  and

$$SNR_{c, ideal} = \frac{P_{c1}L}{N_{01}B} \quad (88)$$

The combining degradation in decibels for IF carrier arraying becomes

$$D_{ifc} = 10 \log_{10} \left[ \frac{L + 2(L-1)e^{-(\sigma_{\Delta\phi, c}^2)/2} + (L-2)(L-1)e^{-\sigma_{\Delta\phi, c}^2}}{L^2} \right] \quad (89)$$

## E. Arraying Combinations

Besides the individual arraying schemes described in this article, combinations of them can be implemented. In particular, SSC can be enhanced with SA and with CA. Similarly, BC can be enhanced with SA and with CA. FSC uses only one set of receiver, subcarrier, and symbol tracking loops, but, again, the performance of the receiver can be improved with SA. A general symbol SNR degradation function, which is applicable to any arraying scheme, is given in Appendix C, Table C-2.

A comparison of all schemes and arraying combinations is depicted in Figs. 15(a) and (b), where the degradations of BC, SSC, FSC, SSC/SA/CA, FSC/SA, BC/SA/CA, SSC/CA, SSC/SA, BC/SA, and BC/CA are all computed versus  $P_T/N_0$  for a fixed  $\Delta = 65.9$  deg. These curves were computed assuming  $B_r = 0.1$  mHz and  $B_n = 135$  kHz for the RTC,  $T/B = 0.0008$  sec<sup>2</sup> for FSC,  $T/B = 0.075$  sec<sup>2</sup> for CA (assumed at IF), and a symbol rate of 34 symbols per sec (sps). From Fig. 15(a), it seems that the three schemes with the least degradation at 20 dB-Hz are FSC/SA, BC/CA/SA, and SSC/SA/CA. Most schemes seem to maintain an 8-dB minimum carrier loop SNR for  $P_T/N_0$  as low as 20 dB-Hz, except for SSC and BC which

lose lock at roughly 24 dB-Hz and BC/CA, and SSC/CA which requires a  $P_T/N_0 \geq 21$  dB-Hz. Recall that the delay adjustment in FSC and FSC/SA was assumed perfect resulting in no degradation. More realistically, a 0.05-dB or a 0.1-dB degradation should be added and, hence, FSC/SA and BC/SA/CA seem to provide identical degradations.

For this particular case, FSC requires 216 sec of integration length (for  $T/B = 0.0008$  sec<sup>2</sup> and  $B = 2 \times 135$  kHz), a rather unrealistic parameter. For a shorter integration time (on the order of a few seconds), the correlator SNR degrades significantly and the differential phase cannot be estimated. The bandwidth  $B$  can be reduced to pass only the first harmonic of the subcarrier, but that still results in unrealistic integration times. The signal can be passed through a "matched filter," which passes the subcarrier harmonics and the data modulation, but rejects the spectrum between the harmonics. The effective bandwidth of such a filter would be on the order of the symbol rate and, hence, results in shorter integration times as long as the subcarrier frequency is a "large" multiple of the symbol rate ( $m \gg 1$ ). The drawback of such a filter is that it is too specific to the signal of interest and needs to be modified for each mission. Moreover, it might

require frequency tuning to center the signal in the band of interest. Another technique to reduce the bandwidth is to only correlate the residual carrier components in order to further shorten the integration time. This is precisely the technique employed in carrier arraying, when implemented at IF. It should be pointed out that even though the phase is adjusted at IF, it can and should be estimated at baseband by mixing the received IF from each antenna with a Doppler and a Doppler rate predict of the signal. The outputs of the mixers consist of a tone with a very low frequency component which requires a "very small" bandwidth  $B$  prior to the correlation. With  $T/B = 0.075 \text{ sec}^2$  and  $T = 3 \text{ sec}$ ,  $B = 2 \times 20 \text{ Hz}$ , which requires the frequency predicts to be correct to within  $\pm 20 \text{ Hz}$ . Even if the error is larger than  $\pm 20 \text{ Hz}$ , a fast Fourier transform can be used to reduce the frequency error at the output of the mixers such that it lies well within  $B/2 \text{ Hz}$ .

As seen from the above example, FSC/SA and BC/SA/CA provide the least degradation and, hence, the "best" performance overall, but BC/CA/SA accomplishes that with reasonable integration times. SA is enhancing the performance in both cases because the carrier component is so weak due to the high modulation index and relatively low received power. For signals with stronger carriers, FSC and FSC/SA would provide similar degradations for all practical purposes, as would BC/CA and BC/SA/CA. It is worth noting at this point that FSC, as presented in this article, compensated for the signal delays up front and then adjusted for the phases. This is the classical arraying performed in radiometry. However in BC/CA, CA is first employed to lock on the signal (hence, a phase adjustment) and later, delay compensation is performed in the BBA to coherently add the data. The latter, which is equivalent in performance to FSC (but with shorter integration times), seems to be favored more by communication engineers whereas FSC seems to be favored more by astrophysicists. The major difference between FSC and BC/CA is the integration length required to estimate the differential phase. BC/CA offers a significant advantage by requiring much shorter integration times for spacecraft with very weak signals and a large subcarrier-to-data-rate ratio.

In either FSC or BC/CA, atmospheric effects can be significant, especially at higher frequencies and in the presence of thunderstorms. Figure 16 depicts the relative phase along baseline "1-3" in the VLA on a clear night and in the presence of thunderstorms. In the latter case, the integration time  $T$  needs to be short to track the phase variation. The resulting combining degradation can be 0.2 dB or even more depending on the scenario.

## IV. Numerical Examples

The results derived in this article were applied to several existing deep-space missions in order to illustrate the differences in combined symbol SNR performance. The missions considered were Pioneer 10, Voyager 2, and Magellan, reflecting weak, medium, and strong signals, respectively, in the DSN. As expected, the weaker the signal, the harder it is to array the antennas.

### A. Pioneer 10

The signal received from Pioneer 10 represents the weakest signal in the DSN. It is an S-band signal with the following characteristics as of May 1990:

- (1) symbol rate  $R_s = 32 \text{ sps}$
- (2) subcarrier frequency  $f_{sc} = 32768 \text{ Hz}$
- (3) modulation index  $\Delta = 65.9 \text{ deg}$

The receiver is assumed to operate with the following parameters:

- (1)  $B_c = 1.5 \text{ Hz}$  for the carrier bandwidth (Block IV Receiver)
- (2)  $B_{sc} = B_{sy} = 0.1 \text{ Hz}$  for subcarrier and symbol tracking loops
- (3)  $B_r = 0.1 \text{ mHz}$  and  $B_n = 135 \text{ kHz}$  for the RTC
- (4)  $T/B = 0.075 \text{ sec}^2$ ,  $B = 2 \times 20 \text{ Hz}$ , and  $T = 3 \text{ sec}$  for carrier arraying
- (5)  $T/B = 0.0008 \text{ sec}^2$  for FSC with regular IF filters ( $B = 2 \times 135 \text{ kHz}$  and  $T = 216 \text{ sec}$ )
- (6)  $T/B = 0.0008 \text{ sec}^2$  for FSC with "matched" filter [ $B = 5 \times (2 \times 50) \text{ Hz}$  and  $T = 0.4 \text{ sec}$ ], where the factor 5 accounts for the first five odd subharmonics of the square-wave subcarrier.

Two array configurations are considered: a 70-m and 34-m STD antenna array, which can provide 0.68-dB gain (over the 70-m antenna) in the ideal case, and an array of two 70-m antennas (providing an ideal 3-dB gain). The degradations for both arrays are shown in Tables 1 and 2, respectively. The 20-dB-Hz signal represents the approximate level at the master antenna, in this case, the 70-m antenna.

In the first array (the 70-m plus STD 34-m antennas), BC and SSC cannot operate due to the inability of the STD 34-m antenna to maintain carrier lock. However, BC/SA and SSC/SA can operate with an 8-dB loop SNR, which is the minimum required to avoid cycle slipping. FSC/SA

achieves the highest loop SNR at 18.2 dB, followed by BC/SA/CA and SSC/SA/CA at 17.7 dB, and followed finally by BC/SA, SSC/SA, and FSC at 11 dB. The smallest degradations are obtained with FSC/SA and BC/SA/CA at about 0.53 dB. Note that the combining loss of FSC at 0.19 dB can be reduced by integrating over longer periods. In the array of two 70-m antennas, all schemes maintain lock as expected with the smallest degradation achieved by FSC/SA at 0.34 dB and the largest achieved by BC at 0.81 dB. FSC/SA seems to be the best arraying scheme for Pioneer 10 and the sideband aiding is essential in reducing the degradation. However, recall that the long integration time required in FSC/SA renders the scheme impractical; hence, BC/CA/SA is really the best scheme for Pioneer 10.

## B. Voyager 2

Unlike Pioneer 10, Voyager 2 can be tracked by all 34-m antennas in the DSN. It represents a medium signal in both received power and data rate. The X-band signal processes the following characteristics:

- (1) symbol rate  $R_s = 43.2$  sps
- (2) subcarrier frequency  $f_{sc} = 360$  kHz
- (3) modulation index  $\Delta = 77$  deg

The receivers are assumed to operate with the following parameters:

- (1)  $B_c = 10$  Hz for the carrier bandwidth
- (2)  $B_{sc} = B_{sy} = 1.0$  Hz for subcarrier and symbol tracking loops
- (3)  $B_r = 1$  mHz and  $B_n = 3.2$  MHz for the RTC
- (4)  $T/B = 0.075 \text{ sec}^2$  for carrier arraying
- (5)  $T/B = 2.0 \times 10^{-7} \text{ sec}^2$ ,  $B = 3.2$  MHz, and  $T = 1.3$  sec for FSC

Table 3 provides the degradations for all arraying schemes for a three-element array of one HEF 34-m and two STD 34-m antennas. This array can provide an ideal 3-dB gain over the HEF 34-m master antenna, with  $P_T/N_0 = 39$  dB-Hz. The second array, whose performance is shown in Table 4, also consists of three elements: one 70-m, one STD 34-m, and one HEF 34-m antenna. The master in this case is the 70-m antenna with  $P_T/N_0 = 45$  dB-Hz. This array can provide a maximum gain of 1.43 dB. BC/SA, BC/CA, and BC/SA/CA can provide

the smallest degradations if the combining loss is maintained below 0.01 dB. On the other hand, FSC/SA provides a better performance for a more realistic 0.07-dB IF degradation. For all practical purposes, both FSC and BC/CA perform equally with realistic integration times.

## C. Magellan

The highest data rate signal is transmitted by Magellan at X-band with

- (1) symbol rate  $R_s = 537.6$  ksp/s
- (2) subcarrier frequency  $f_{sc} = 960$  kHz
- (3) modulation index  $\Delta = 78$  deg

Tables 5 and 6 provide the degradations for an array of one HEF 34-m and one STD 34-m antenna (providing a 1.76-dB ideal gain over the HEF 34-m master antenna) and another array of one 70-m, one HEF 34-m, and one STD 34-m antenna (providing a 1.43-dB ideal gain over the 70-m master antenna). The receivers are assumed to operate with

- (1)  $B_c = 30$  Hz for carrier bandwidth
- (2)  $B_{sc} = B_{sy} = 3.0$  Hz for subcarrier and symbol tracking loops
- (3)  $B_r = 10$  mHz and  $B_n = 4.5$  MHz for the RTC
- (4)  $T/B = 0.075 \text{ sec}^2$  for carrier arraying
- (5)  $T/B = 1.0 \times 10^{-10} \text{ sec}^2$  for FSC

In this case, all combining methods provide near-optimum performances for both arrays.

## V. Conclusions

Four different arraying schemes have been investigated and these include symbol stream combining, baseband combining, carrier arraying, and full spectrum combining. For DSN applications where telemetry signal reception is the primary concern, BC/CA and BC/CA/SA provide the best arraying schemes for very weak signals with large subcarrier-frequency-to-data-rate ratios. FSC and FSC/SA are not well suited for these scenarios, but can be made so by employing "matched filters" at the cost of additional complexity. For moderate to high signal levels, FSC and BC/CA are both well suited and provide comparable performances.

## Acknowledgments

The authors thank D. Rogstad for his numerous discussions and J. Ulvestad for providing Fig. 16 on the VLA thunderstorm data. Special thanks to G. Resch for suggesting the idea of the article and to J. Statman for his valuable comments while reviewing this article.

## References

- [1] J. W. Layland, P. J. Napier, and A. R. Thompson, "A VLA Experiment—Planning for Voyager at Neptune," *TDA Progress Report 42-82*, vol. April–June 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 136–142, August 15, 1985.
- [2] J. S. Ulvestad, "Phasing the Antennas of the Very Large Array for Reception of Telemetry from Voyager 2 at Neptune Encounter," *TDA Progress Report 42-94*, vol. April–June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 257–273, August 15, 1988.
- [3] J. Yuen, *Deep Space Telecommunications Systems Engineering*, New York: Plenum Press, 1982.
- [4] S. Hinedi, "A Functional Description of the Advanced Receiver," *TDA Progress Report 42-100*, vol. October–December 1989, Jet Propulsion Laboratory, Pasadena, California, February 15, 1990, pp. 131–149.
- [5] R. Sfeir, S. Aguirre, and W. Hurd, "Coherent Digital Demodulation of a Residual Signal Using IF Sampling," *TDA Progress Report 42-78*, vol. April–June 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 135–142, August 15, 1984.
- [6] W. Lindsey and M. Simon, *Telecommunication Systems Engineering*, New Jersey: Prentice-Hall, 1973.
- [7] W. Hurd and S. Aguirre, "A Method to Dramatically Improve Subcarrier Tracking," *TDA Progress Report 42-86*, vol. April–June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 103–110, August 15, 1986.
- [8] W. J. Hurd, J. Rabkin, M. D. Russel, B. Siev, H. W. Cooper, T. O. Anderson, and P. U. Winter, "Antenna Arraying of Voyager Telemetry Signals by Symbol Stream Combining," *TDA Progress Report 42-86*, vol. April–June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 131–142, August 15, 1986.
- [9] D. Divsalar, "Symbol Stream Combining Versus Baseband Combining for Telemetry Arraying," *TDA Progress Report 42-74*, vol. April–June 1983, Jet Propulsion Laboratory, Pasadena, California, pp. 13–28, August 15, 1983.
- [10] M. K. Simon and A. Mileant, *Performance Analysis of the DSN Baseband Assembly Real-Time Combiner*, JPL Publication 84-94, Revision 1, Jet Propulsion Laboratory, Pasadena, California, May 1, 1985.
- [11] R. M. Hjellming, ed., *An Introduction to the NRAO Very Large Array*, National Radio Astronomy Observatory, Socorro, New Mexico, 1980.
- [12] D. Divsalar and J. Yuen, "Improved Carrier Tracking Performance with Coupled Phase-Locked Loops," *TDA Progress Report 42-66*, vol. April–June 1981, Jet Propulsion Laboratory, Pasadena, California, pp. 148–171, August 15, 1981.

**Table 1. Pioneer 10, one 70-m and one STD 34-m antenna array**

Arraying scheme	$P_T/N_0$ , dB-Hz	Total degradation, dB	Carrier degradation, dB	Subcarrier degradation, dB	Symbol degradation, dB	RTC or IF degradation, dB
BC	20.00	no	carrier	lock	—	—
BC+SA	20.00	-0.614	-0.17	-0.25	-0.07	-0.12
BC+CA	20.00	-0.792	-0.34	-0.26	-0.08	-0.12
BC+SA+CA	20.00	-0.526	-0.07	-0.26	-0.08	-0.12
SSC	20.00	no	carrier	lock	—	—
SSC+SA	20.00	-0.670	-0.17	-0.39	-0.11	0.00
SSC+CA	20.00	-0.849	-0.34	-0.40	-0.11	0.00
SSC+SA+CA	20.00	-0.583	-0.07	-0.40	-0.11	0.00
FSC	20.00	-0.874	-0.35	-0.26	-0.08	-0.19
FSC+SA	20.00	-0.593	-0.07	-0.26	-0.08	-0.19
FSC (matched filter)	20.00	-0.874	-0.35	-0.26	-0.08	-0.19
FSC+SA (matched filter)	20.00	-0.593	-0.07	-0.26	-0.08	-0.19

**Table 2. Pioneer 10, two 70-m antenna arrays**

Arraying scheme	$P_T/N_0$ , dB-Hz	Total degradation, dB	Carrier degradation, dB	Subcarrier degradation, dB	Symbol degradation, dB	RTC or IF degradation, dB
BC	20.00	-0.812	-0.40	-0.19	-0.06	-0.17
BC+SA	20.00	-0.487	-0.08	-0.19	-0.06	-0.17
BC+CA	20.00	-0.608	-0.20	-0.19	-0.06	-0.17
BC+SA+CA	20.00	-0.475	-0.06	-0.19	-0.06	-0.17
SSC	20.00	-0.768	-0.40	-0.29	-0.08	0.00
SSC+SA	20.00	-0.444	-0.08	-0.29	-0.08	0.00
SSC+CA	20.00	-0.565	-0.20	-0.29	-0.08	0.00
SSC+SA+CA	20.00	-0.432	-0.06	-0.29	-0.08	0.00
FSC	20.00	-0.509	-0.20	-0.19	-0.06	-0.07
FSC+SA	20.00	-0.347	-0.04	-0.19	-0.06	-0.07
FSC (matched filter)	20.00	-0.509	-0.20	-0.19	-0.06	-0.07
FSC+SA (matched filter)	20.00	-0.347	-0.04	-0.19	-0.06	-0.07

**Table 3. Voyager 2, one master HEF 34-m and two STD 34-m antenna arrays**

Arraying scheme	$P_T/N_0$ , dB-Hz	Total degradation, dB	Carrier degradation, dB	Subcarrier degradation, dB	Symbol degradation, dB	RTC or IF degradation, dB
BC	39.00	-0.346	-0.16	-0.11	-0.03	-0.04
BC+SA	39.00	-0.219	-0.04	-0.11	-0.03	-0.04
BC+CA	39.00	-0.236	-0.05	-0.11	-0.03	-0.04
BC+SA+CA	39.00	-0.197	-0.02	-0.11	-0.03	-0.04
SSC	39.00	-0.548	-0.16	-0.31	-0.08	0.00
SSC+SA	39.00	-0.422	-0.04	-0.31	-0.08	0.00
SSC+CA	39.00	-0.439	-0.05	-0.31	-0.08	0.00
SSC+SA+CA	39.00	-0.400	-0.02	-0.31	-0.08	0.00
FSC	39.00	-0.284	-0.06	-0.11	-0.03	-0.09
FSC+SA	39.00	-0.235	-0.01	-0.11	-0.03	-0.09

**Table 4. Voyager 2, one 70-m, one STD 34-m, and one HEF 34-m antenna array**

Arraying scheme	$P_T/N_0$ , dB-Hz	Total degradation, dB	Carrier degradation, dB	Subcarrier degradation, dB	Symbol degradation, dB	RTC or IF degradation, dB
BC	45.00	-0.130	-0.06	-0.05	-0.01	-0.01
BC+SA	45.00	-0.084	-0.01	-0.05	-0.01	-0.01
BC+CA	45.00	-0.091	-0.02	-0.05	-0.01	-0.01
BC+SA+CA	45.00	-0.077	-0.01	-0.05	-0.01	-0.01
SSC	45.00	-0.208	-0.06	-0.12	-0.03	0.00
SSC+SA	45.00	-0.163	-0.01	-0.12	-0.03	0.00
SSC+CA	45.00	-0.170	-0.02	-0.12	-0.03	0.00
SSC+SA+CA	45.00	-0.156	-0.01	-0.12	-0.03	0.00
FSC	45.00	-0.148	-0.02	-0.05	-0.01	-0.07
FSC+SA	45.00	-0.134	-0.01	-0.05	-0.01	-0.07

**Table 5. Magellan, one master HEF 34-m and one STD 34-m antenna array**

Arraying scheme	$P_T/N_0$ , dB-Hz	Total degradation, dB	Carrier degradation, dB	Subcarrier degradation, dB	Symbol degradation, dB	RTC or IF degradation, dB
BC	59.00	-0.022	-0.01	-0.01	0.00	0.00
BC+SA	59.00	-0.022	-0.01	-0.01	0.00	0.00
BC+CA	59.00	-0.022	-0.01	-0.01	0.00	0.00
BC+SA+CA	59.00	-0.022	-0.01	-0.01	0.00	0.00
SSC	59.00	-0.027	-0.01	-0.02	0.00	0.00
SSC+SA	59.00	-0.027	-0.01	-0.02	0.00	0.00
SSC+CA	59.00	-0.027	-0.01	-0.02	0.00	0.00
SSC+SA+CA	59.00	-0.027	-0.01	-0.02	0.00	0.00
FSC	59.00	-0.036	-0.01	-0.01	0.00	-0.02
FSC+SA	59.00	-0.036	-0.01	-0.01	0.00	-0.02

**Table 6. Magellan, one 70-m, one HEF 34-m, and one STD 34-m antenna array**

Arraying scheme	$P_T/N_0$ , dB-Hz	Total degradation, dB	Carrier degradation, dB	Subcarrier degradation, dB	Symbol degradation, dB	RTC or IF degradation, dB
BC	65.00	-0.015	-0.01	-0.01	0.00	0.00
BC+SA	65.00	-0.015	-0.01	-0.01	0.00	0.00
BC+CA	65.00	-0.015	-0.01	-0.01	0.00	0.00
BC+SA+CA	65.00	-0.015	-0.01	-0.01	0.00	0.00
SSC	65.00	-0.021	-0.01	-0.01	0.00	0.00
SSC+SA	65.00	-0.021	-0.01	-0.01	0.00	0.00
SSC+CA	65.00	-0.021	-0.01	-0.01	0.00	0.00
SSC+SA+CA	65.00	-0.021	-0.01	-0.01	0.00	0.00
FSC	65.00	-0.031	-0.01	-0.01	0.00	-0.02
FSC+SA	65.00	-0.031	-0.01	-0.01	0.00	-0.02

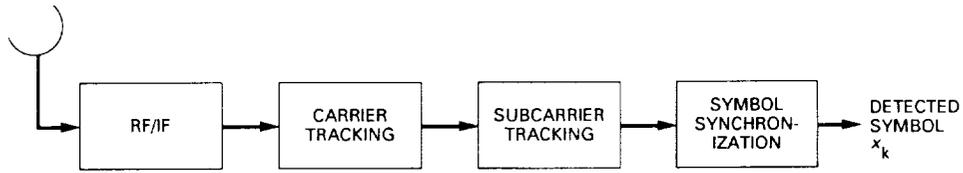


Fig. 1. A general coherent receiver model.

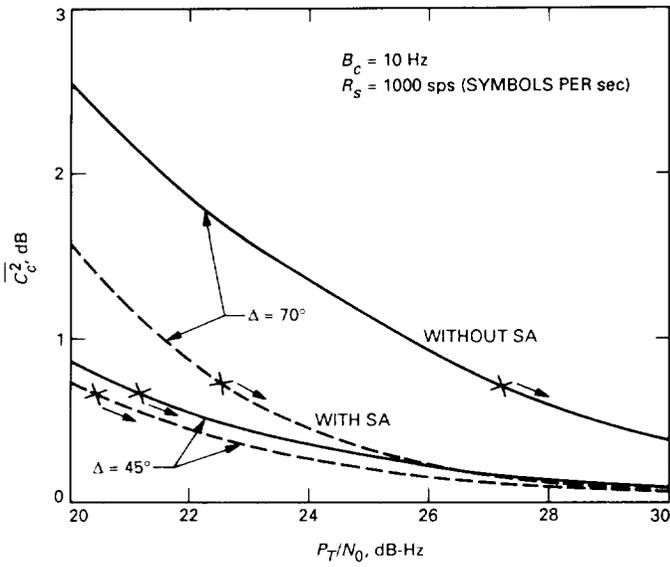


Fig. 2. Symbol SNR degradation due to imperfect carrier reference.

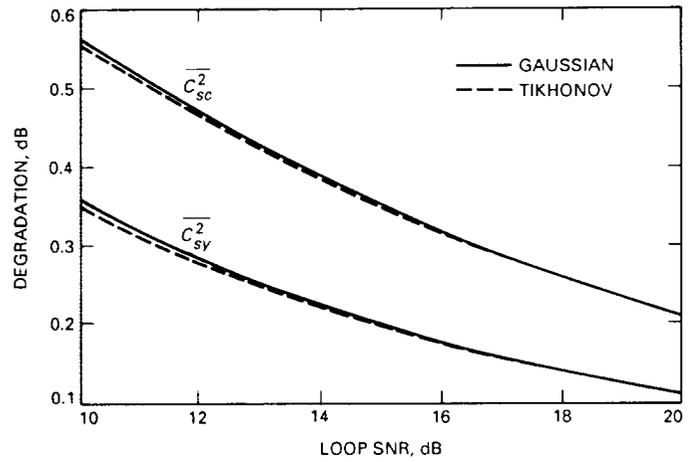


Fig. 3. Symbol SNR degradation in the presence of subcarrier and symbol phase jitter.

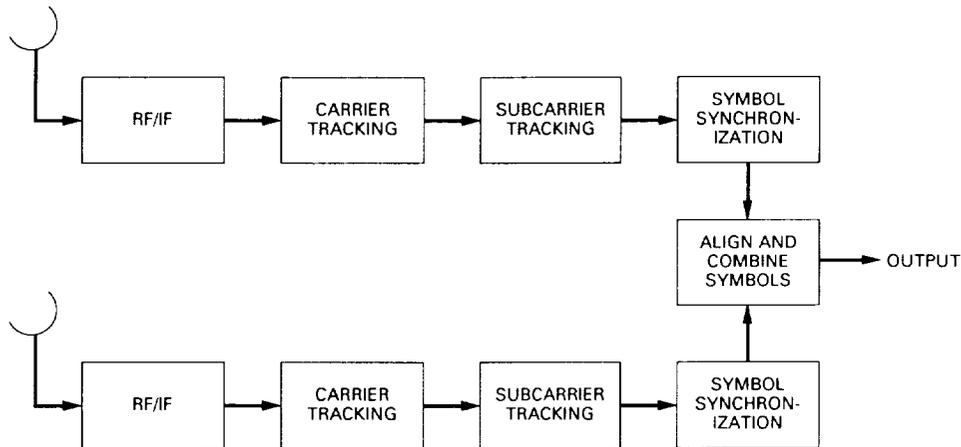


Fig. 4. Symbol stream combining.

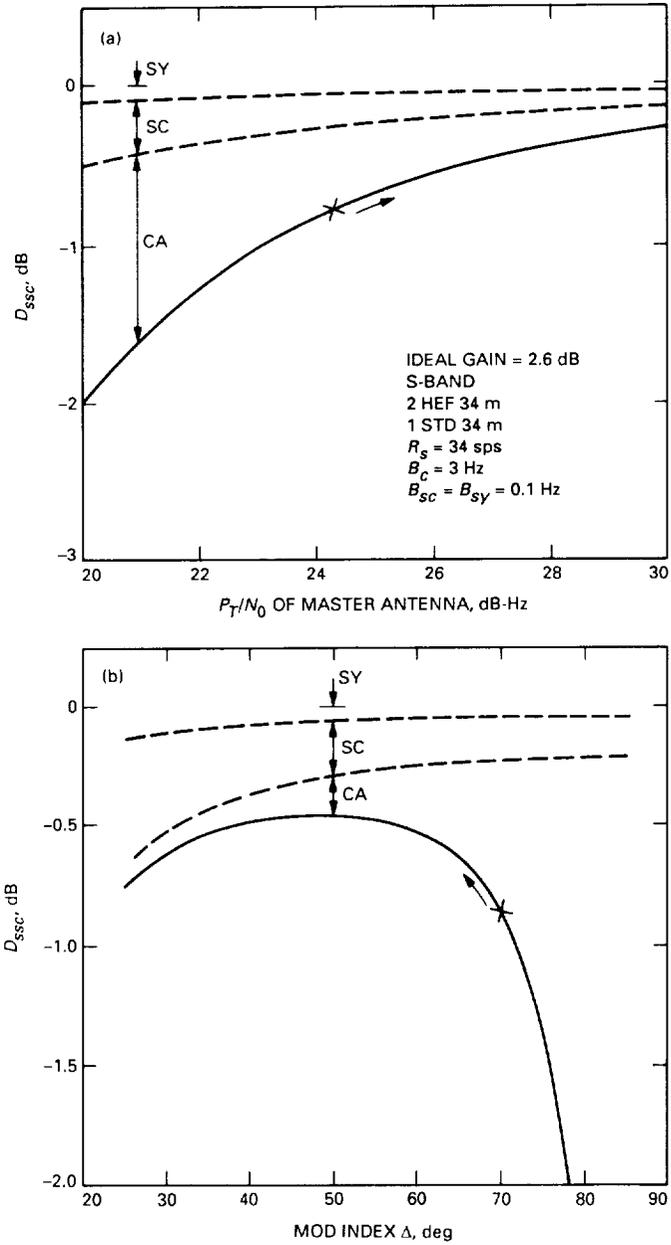


Fig. 5. The  $D_{SSC}$  versus: (a)  $P_T/N_0$  for  $\Delta = 65.9$  deg, and  
 (b)  $\Delta$  for  $P_T/N_0 = 25$  dB-Hz.

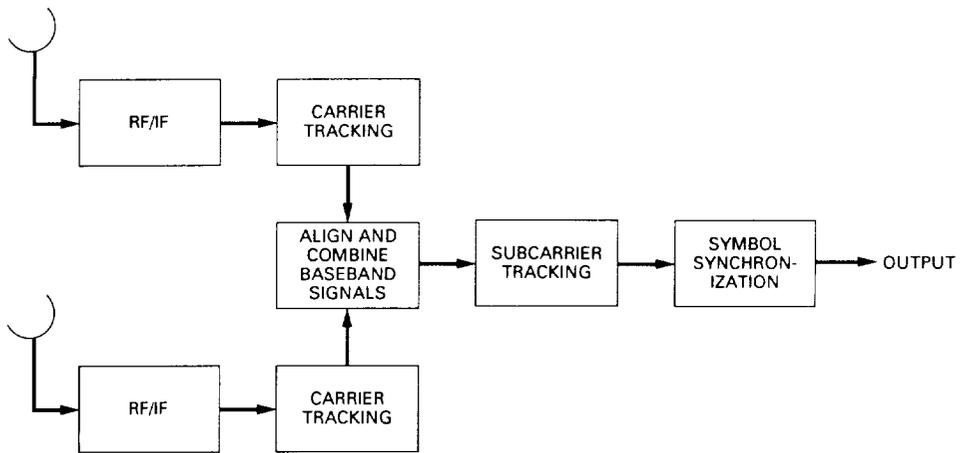


Fig. 6. Baseband combining.

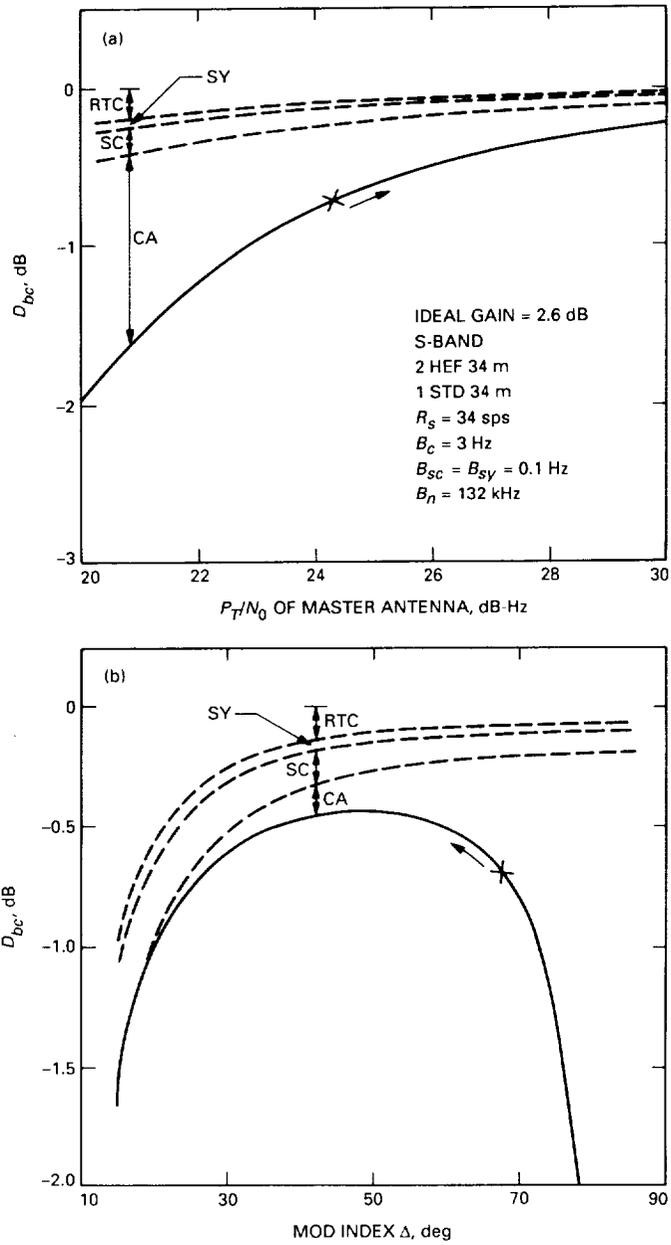


Fig. 7. The  $D_{bc}$  versus: (a)  $P_T/N_0$  for  $\Delta = 65.9$  deg, and (b)  $\Delta$  for  $P_T/N_0 = 25$  dB-Hz.

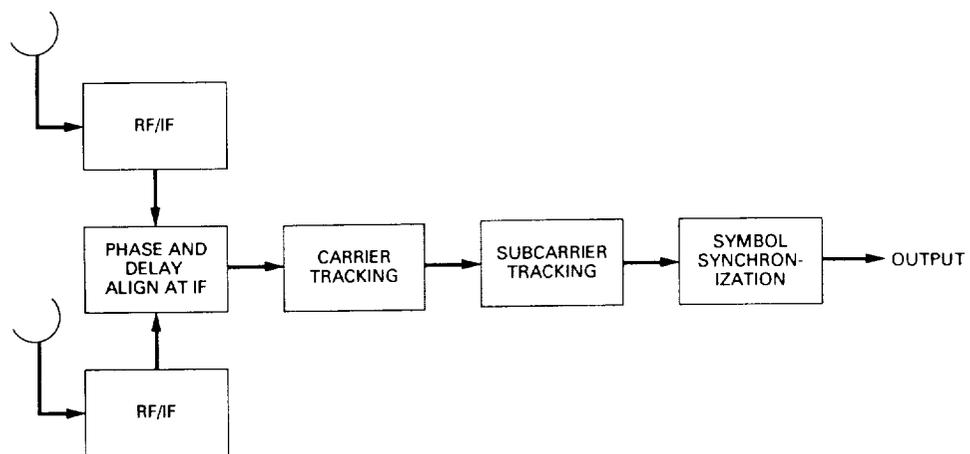


Fig. 8. Full spectrum combining.

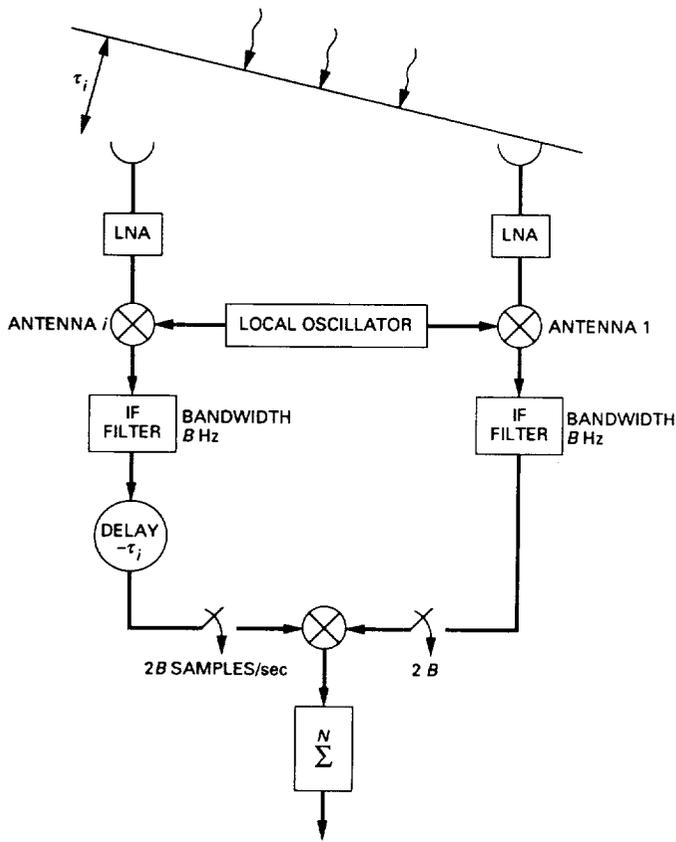


Fig. 9. An interferometric antenna pair.

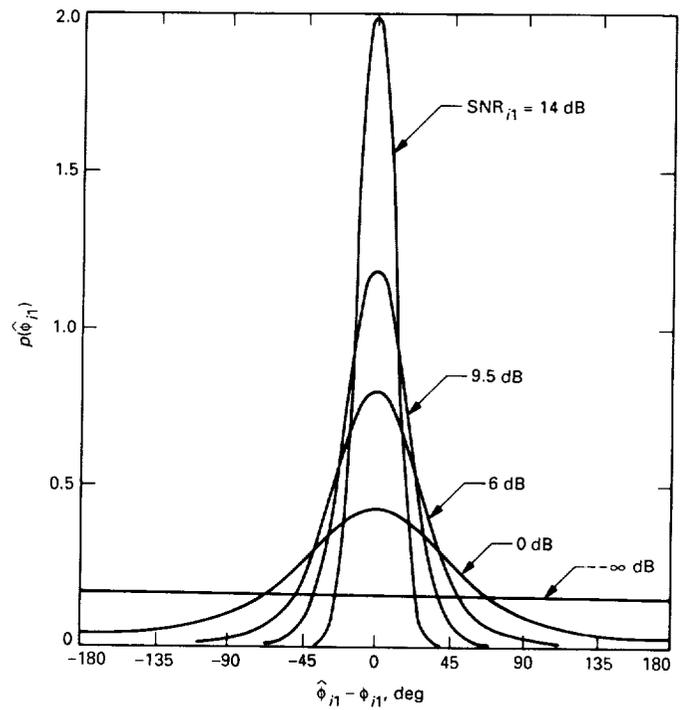


Fig. 10. The probability distribution of measured phase as a function of  $\hat{\phi} - \phi$  for a number of signal-to-noise ratios.

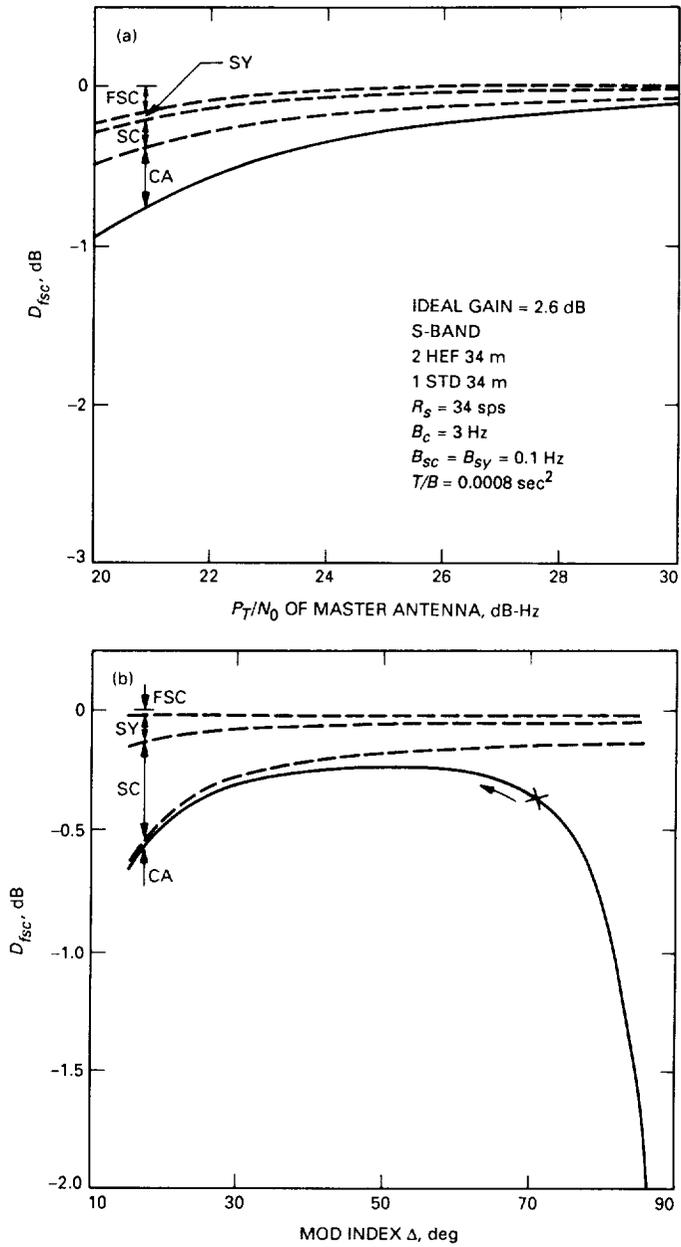


Fig. 11. The  $D_{fsc}$  versus: (a)  $P_T/N_0$  for  $\Delta = 65.9$  deg, and (b)  $\Delta$  for  $P_T/N_0 = 25$  dB-Hz.

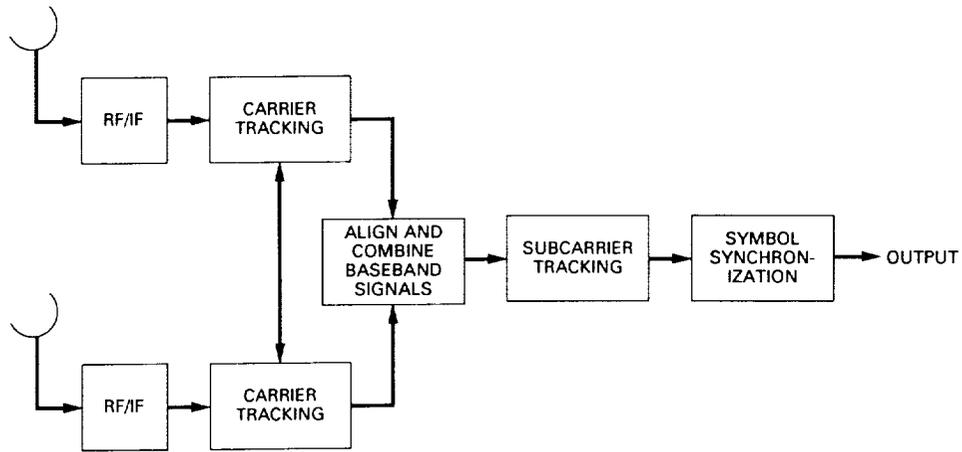


Fig. 12. Carrier arraying with baseband combining (CA/BC).

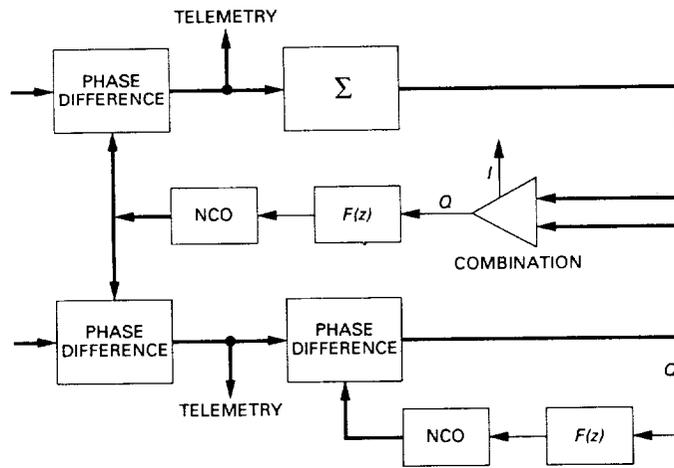


Fig. 13. A baseband implementation of carrier arraying.

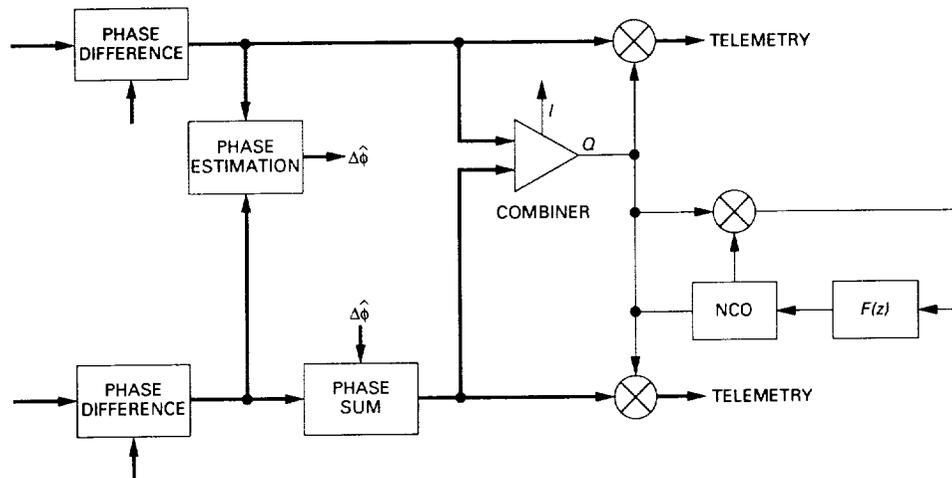


Fig. 14. An IF implementation of carrier arraying.

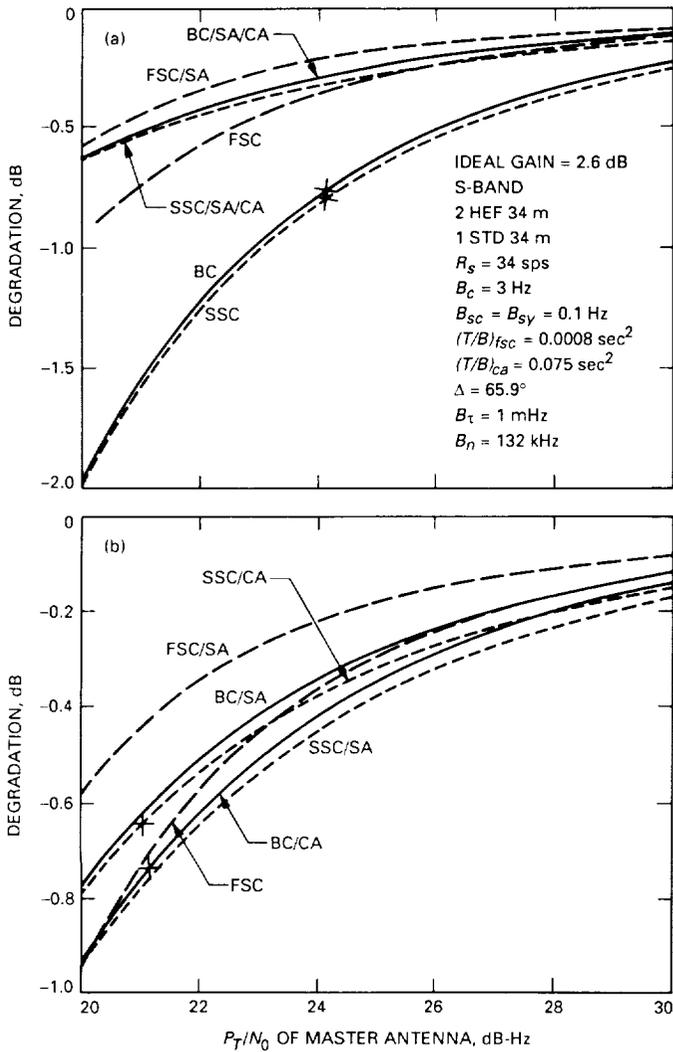


Fig. 15. Comparison of SSC, FSC, and FSC/SA with: (a) BC, SSC/SA/CA, and BC/SA/CA, and (b) BC/CA, BC/SA, and SSC/CA.

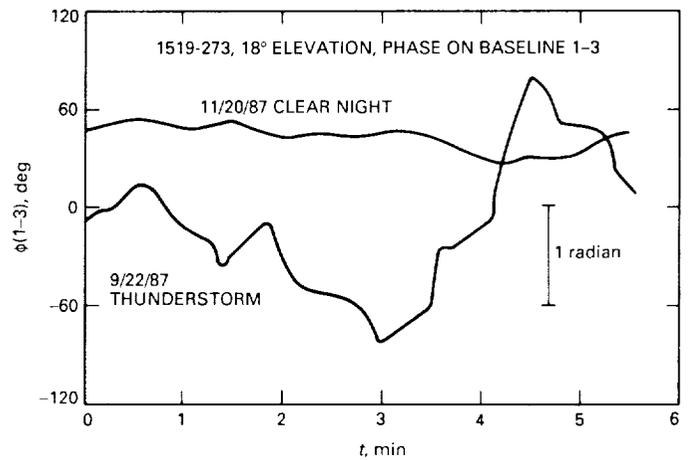


Fig. 16. Very Large Array thunderstorm data at 8.4 GHz.

## Appendix A

### Gamma Factors for DSN Antennas

Table A-1 summarizes the  $\gamma_i$  factors,<sup>1</sup> defined by Eq. (23), for several DSN antennas at both S-band (2.2 to 2.3 GHz) and X-band (8.4 to 8.5 GHz). Conceptually, these gamma factors represent the antenna gain/noise temperature ratios normalized by the gain/noise temper-

ature of the largest antenna. Here HEF denotes high-efficiency antenna and STD a standard antenna.

The numbers presented below should be used in a relative sense, not in an absolute sense. For example, for a three-element array consisting of one HEF 34-m antenna and two STD 34-m antennas at S-band, the master antenna (in this case, the HEF 34 m) will have  $\gamma_1 = 1$  and the other two antennas would have  $\gamma_2 = \gamma_3 = 0.13/0.26 = 0.5$ .

<sup>1</sup> *Deep Space Network/Flight Project Interface Design Handbook*, Document 810-5, Rev. D, Vol. I (internal document), Jet Propulsion Laboratory, Pasadena, California, Modules TCI-10, TCI-30, and TLM-10, 1988.

**Table A-1. Gamma factors for DSN antennas**

Antenna size	Frequency band	$\gamma_i$
70 m	S-band	1.00
34 m STD	S-band	0.17
34 m HEF	S-band	0.07
70 m	X-band	1.00
34 m STD	X-band	0.13
34 m HEF	X-band	0.26

## Appendix B

### Closed-Loop Performance

Typically, one would like to limit the IF combining losses expressed by Eq. (75) to some prespecified maximum value, say  $D_{max}$ . Solving Eq. (75) for  $\sigma_{\Delta\phi_{max}}^2$  yields

$$\sigma_{\Delta\phi_{max}}^2 < -2 \ln \left[ \frac{10^{(D_{max}/10)} L - 1}{L - 1} \right] \quad (\text{B-1})$$

The variance of the phase estimate,  $\hat{\phi}_{i1}$ , can be reduced by either increasing the correlation time  $T$  in Eq. (57) or by tracking the phase error process in a closed-loop fashion. Note that the value of  $B$  in Eq. (57) is set by the bandwidth of the telemetry spectrum and cannot be reduced at will.

In the simplest closed-loop implementation of the full spectrum combining scheme, phase-error estimates can be updated using the following difference equation:

$$\hat{\theta}(n) = \hat{\theta}(n-1) + \alpha\phi(n) \quad (\text{B-2})$$

where the value of  $\alpha$  can be set between 0.2 and 0.5 and  $\hat{\theta}(n)$  is the filtered phase error estimate. The above difference equation gives the following loop transfer function:

$$G(z) = \frac{\hat{\Theta}(z)}{\Phi(z)} = \frac{\alpha}{z-1} \quad (\text{B-3})$$

The variance of the closed-loop phase error process will now be

$$\sigma_{i1}^2 = I_1 \sigma_{\Delta\phi, i1}^2 = I_1 \frac{N_{o1} N_{oi} B}{P_1 P_i 2T} \quad (\text{B-4})$$

where

$$I_1 = \frac{1}{2\pi j} \oint |H(z)|^2 \frac{dz}{z} \quad (\text{B-5})$$

and  $H(z) = G(z)/[1 + G(z)]$ . Using the above  $G(z)$  gives

$$I_1 = \frac{\alpha}{2 - \alpha} \quad (\text{B-6})$$

As an example, for  $\alpha = 0.2$ ,  $I_1 = 0.11$  and the variance of the phase jitter is reduced by a factor of 10.

## Appendix C

### Generalization of the Symbol SNR Degradation Function

By comparing the symbol SNR degradation factors that were obtained for different arraying schemes, one arrives at the following general equation of SNR degradation for combinations of arraying schemes:

$$D = 10 \log_{10} \left( \frac{\overline{C_A^2}}{\frac{\sum_{i=1}^L \gamma_i^2 \overline{C_{B_i}^2} + \sum_{\substack{i,j \\ i \neq j}} \gamma_i \gamma_j \overline{C_{B_i} C_{B_j}}}{\Gamma^2}} \right) \quad (\text{C-1})$$

where the particular signal reduction factors  $C_A$  and  $C_{B_i}$  are summarized in Table C-1.

Without SA, the carrier loop SNR is  $\rho_{c,r} \triangleq 1/\sigma_{c,r}^2 = P_C/(N_0 B_c)$ , while with SA, the loop SNR becomes  $\rho_c = \rho_{c,r} + \rho_{c,s}$ , where  $\rho_{c,s}$  is given in Eq. (10). Note that in BC and FSC,  $P_D$  is the combined data power, reduced somewhat by the combining loss.

**Table C-1. Comparison of signal reduction factors for different arraying schemes**

Item	SSC	SSC+CA	BC	BC+CA	FSC
$C_A$	1	$C_c$	$C_{sc} C_{sy}$	$C_c C_{sc} C_{sy}$	$C_c C_{sc} C_{sy}$
$C_{B_i}$	$C_{ci} C_{sci} C_{syi}$	$C_{sci} C_{syi}$	$C_{ci} C_{\tau i}$	$C_{\tau i}$	$C_{IF_i}$

Table C-2. Signal reduction factors

Definition	Reduction function	First moment	Second moment	Variance
$C_c$	$\cos \phi$	$\frac{I_1(\rho_c)}{I_0(\rho_c)}$	$\frac{1}{2} \left[ 1 + \frac{I_2(\rho_c)}{I_0(\rho_c)} \right]$	$\sigma_{c,r}^2 = \frac{N_0 B_c}{P_C}$
$C_{sc}$	$1 - \frac{2}{\pi}  \phi_{sc} $	$1 - \left( \frac{2}{\pi} \right)^{3/2} \sigma_{sc}$	$1 - \sqrt{\frac{32}{\pi^3}} \sigma_{sc} + \frac{4}{\pi^2} \sigma_{sc}^2$	$\left( \frac{\pi}{2} \right)^2 \frac{B_{sc} W_{sc}}{R_s \left( \frac{E_s}{N_0} \right)^2} \left( 1 + \frac{E_s}{N_0} \right)$
$C_{sy}$	$1 - \frac{1}{2\pi}  \phi_{sy} $	$1 - \sqrt{\frac{1}{2\pi}} \frac{\sigma_{sy}}{\pi}$	$1 - \sqrt{\frac{2}{\pi}} \frac{\sigma_{sy}}{\pi} + \frac{\sigma_{sy}^2}{4\pi^2}$	$2\pi^2 \frac{B_{sy} W_{sy}}{R_s \frac{E_s}{N_0} \Phi^2 \sqrt{\frac{E_s}{N_0}}}$
$C_r$	$1 - 4m \tau_i $	$1 - 4m\sqrt{\frac{2}{\pi}} \sigma_{\tau i}$	$1 - 8m\sqrt{\frac{2}{\pi}} \sigma_{\tau i} + 16m^2 \sigma_{\tau i}^2$	$\frac{B_{\tau i}}{B_n 32m^2} \left\{ \frac{1}{\left[ \text{erf} \left( \sqrt{\frac{P_{D_i}}{2\sigma_i^2}} \right) \text{erf} \left( \sqrt{\frac{P_{D_i}}{2\sigma_i^2}} \right) \right]^2} - 1 \right\}$
$C_{IF}, C_{IF}^*$	$e^{j[\Delta\phi_{i1}(t_k) - \Delta\phi_{j1}(t_k)]}$	$e^{-\frac{1}{2} [\sigma_{\Delta\phi_{i1}}^2 + \sigma_{\Delta\phi_{j1}}^2]}$	—	$\frac{N_{0i} N_{0j} B}{P_{c1} P_{c2} 2T}$

512-17  
P-17

N91-18320

JJ 574450

# The Advanced Receiver II: Telemetry Test Results in CTA 21

S. Hinedi and R. Bevan  
Communications Systems Research Section

M. Marina  
Radio Frequency and Microwave Subsystems Section

*This article describes telemetry tests with the Advanced Receiver II (ARX II) in Compatibility Test Area 21. The ARX II was operated in parallel with a Block-III Receiver/baseband processor assembly combination (BLK-III/BPA) and a Block-III Receiver/subcarrier demodulation assembly/symbol synchronization assembly combination (BLK-III/SDA/SSA). The telemetry simulator assembly provided the test signal for all three configurations, and the symbol signal-to-noise ratio as well as the symbol error rates were measured and compared. Furthermore, bit-error rates were also measured by the system performance test computer for all three systems. Results indicate that the ARX-II telemetry performance is comparable and sometimes superior to the BLK-III/BPA and BLK-III/SDA/SSA combinations.*

## I. Introduction

The Advanced Receiver (ARX II) [1] is a digital system that performs carrier, subcarrier, and symbol detection, as well as Doppler extraction. The latter function was demonstrated successfully at Goldstone [2] in June 1989. The telemetry functions have been added since then and tested in Compatibility Test Area 21 (CTA21). Specifically, residual carrier, subcarrier, and symbol synchronization have been added, including a symbol signal-to-noise ratio (SSNR) estimator. Sideband aiding, which requires a carrier Costas loop, and quadrature phase-shift keying capabilities have not been added yet.

This article describes the ARX-II telemetry tests which were conducted in CTA21. The ARX II was operated in parallel with a Block-III Receiver/baseband processor assembly combination (BLK-III/BPA) and a Block-III Re-

ceiver/subcarrier demodulation assembly/symbol synchronization assembly combination (BLK-III/SDA/SSA). The BPA is a baseband assembly (BBA) without the real-time combining capability and, hence, is equivalent to the BBA for the purpose of these tests.

## II. Test Objectives

The objectives of the ARX-II telemetry tests at CTA21 were to:

- (1) Demonstrate the added telemetry capability of the ARX II to perform subcarrier demodulation and symbol synchronization at different data rates and at varying signal-to-noise ratios.
- (2) Test the interface between the ARX II and the telemetry processor assembly (TPA).

- (3) Compare the ARX-II SSNR estimator's results with another SSNR estimate determined from measuring the symbol error rate (SER), referred to as the symbol error rate SNR (SERSNR), as a measure of the accuracy of the ARX-II SSNR estimator.
- (4) Compare the telemetry performance of the ARX II to that of the BLK-III/BPA and BLK-III/SDA/SSA combinations.
- (5) Measure telemetry performance with bit and symbol error rates for all three systems.

### III. Test Description

The ARX-II telemetry test configuration in CTA21 is depicted in Fig. 1. The telemetry simulator assembly (TSA) provided a test signal consisting of a semirandom symbol stream modulated on a square wave subcarrier. Internal to the TSA, a pseudorandom sequence is generated to simulate the transmitted bits. These are later coded using a (7,1/2) convolutional code to produce the symbols which modulate the subcarrier. This telemetry baseband signal, consisting of subcarrier and data ( $Sc \times D$ ), forms the input to an exciter which produces an uplink S-band modulated signal at 2113 MHz. A digitally controlled oscillator (DCO) produces a 44-MHz IF and the frequency of this intermediate frequency (IF) signal is multiplied by the integer 48 to obtain the S-band (2112-MHz) uplink. The exciter phase modulator has a specification bandwidth of 2 MHz (which would become a crucial parameter at high data rates) and outputs a constant power level.

The output of the exciter ( $C \times Sc \times D$ ) is then converted to a downlink S-band signal (2295 MHz) by the translator which multiplies the frequency by the ratio 240/221. This signal is then fed into an S-band field-effect-transistor (FET) amplifier with a 4.5-dB noise figure, a 50-dB gain, and a bandwidth from 2 to 3 GHz. The operating system noise temperature was approximately 500 kelvin. The output of the S-band amplifier ( $C \times Sc \times D + N$ ) was then sent simultaneously to the Block-III Receiver and to the multimission receiver (MMR). The MMR downconverted the S-band signal to a 300-MHz IF signal that was sent to the ARX II.

The ARX II was operated in parallel with a BLK-III/BPA and a BLK-III/SDA/SSA combination. SSNR measurements were performed by the ARX II, the BPA, and the SSA using the split-symbol moment estimator [3]. Moreover, both the transmitted symbols and the symbol clock from the TSA were passed to the ARX II, the BPA, and the SSA where SER measurements were performed.

The latter measurements were mapped into an equivalent SSNR, referred to as SERSNR, assuming an additive white Gaussian noise channel. Bit-error rates (BERs) were also measured and compared for all three systems for different data rates. Symbols at the output of the SSA were fed to TPA 1, while the symbols from the output of the ARX II and the BPA shared the same input of the TPA 3 through a telemetry switch. The decoded bits from TPA 1 and TPA 3 were sent to the system performance test (SPT) computer where BER measurements took place.

Telemetry signal path verification tests were run with different test signals to verify the flow of data from the TSA to the TPA. Telemetry tests for different data rates at varying signal-to-noise ratios at the input of the receiver were done according to prepared configuration test tables. These tables represent current flight missions for different spacecraft and cover signals from the Pioneer 10 low data rate (16 bits per sec—bps) through the Magellan high data rate (268.8 kbps) signal.

The test signal was calibrated for different telemetry tests. A carrier suppression procedure was used to set up the modulation index and a  $Y$ -factor procedure was used to set up the bit SNR ( $E_b/N_0$ ) at the input of the receiver. The modulation index adjustment procedure is described in Appendix A, while the  $Y_F$  setup is discussed in Appendix B. Accuracy of both procedures provides a worst-case uncertainty of  $\pm 0.5$  dB in the calibrated test signal SSNR.

### IV. Test Results

The tests performed can be divided into two main classes: general telemetry tests comparing the three configurations and specific tests for the Pioneer 10 spacecraft. The objective of the first set of tests was to compare the three systems on a relative basis, while the results of the latter tests were compared to theoretically predicted results.

It is worth noting at this point the difference between SSNR degradation and SSNR loss. SSNR degradation is defined as the average reduction in SNR at the symbol matched filter output due to imperfect synchronization caused by carrier, subcarrier, or symbol tracking. For example, the SNR degradation due to imperfect carrier reference is given by

$$D_c = \mathcal{E}\{\cos^2 \phi_c\} \quad (1)$$

where  $\mathcal{E}$  denotes expectation over the carrier phase error  $\phi_c$ . SSNR loss, on the other hand, is defined as the ad-

ditional SSNR needed in the presence of imperfect synchronization in order to achieve the same symbol error probability for the case of perfect synchronization. SSNR loss is shown in Fig. 2(a) and is specified at a particular SSNR or the equivalent symbol error rate. SSNR loss is typically larger than the SNR degradation and both were utilized in analyzing the measured data. Bit SNR (BSNR) loss is similarly defined as the additional BSNR required to achieve a certain BER. The performance of the (7,1/2) convolutional code (assuming Viterbi decoding with symbols quantized to three bits), in an additive white Gaussian noise channel, is depicted in Fig. 2(b) and was used to assess the bit-error rate SNR (BERSNR) loss.

### A. Measurement Accuracy

Since the results are based solely on measurements, it is worthwhile at this point to discuss their confidence level. Note that the outcome of each symbol (or bit) detection is either “no symbol error” or “symbol error,” i.e., a binary decision. Let  $\mathcal{X}$  be a random variable denoting the number of “symbol errors” in a test of  $n$  symbols. Then,  $\mathcal{X}$  can be modeled by a binomial distribution with mean  $np$  and variance  $np(1-p)$ , where  $p$  denotes the expected probability of symbol error (i.e., SER). Hence, the error level (the ratio of the standard deviation to the mean) becomes  $\sqrt{(1-p)/np}$ , with an accuracy level of  $1 - \sqrt{(1-p)/np}$ . As an example, when testing a system with an expected probability of symbol error of 2.288 percent (SSNR = 3.0 dB) using 10,000 symbols, the results will be correct with a 93.4 percent accuracy level. On the other hand, for a 0.595 percent SER (SSNR = 5 dB), the accuracy level decreases to 87 percent. To translate the SER accuracy to its counterpart in SSNR, the performance of binary phase-shift key (BPSK) in an additive white Gaussian noise channel is used, Fig. 2(a). For example, the 93.4 percent accuracy at an SSNR of 3 dB translates into an SER of  $0.02288 \pm 6.6$  percent (or a maximum SER of 0.02439). From Fig. 2(a), the latter SER corresponds to 2.85 dB in SSNR and, hence, a deviation or accuracy of  $\pm 0.15$  dB (3 – 2.85 dB) in SSNR. Similar accuracies can be obtained in terms of BSNR using Fig. 2(b).

### B. Relative Performances of ARX II, BLK-III/BPA, and BLK-III/SDA/SSA

Throughout the tests to follow, the data format was non-return-to-zero (NRZ) modulated on a square-wave subcarrier with a  $\pm 0.5$ -dB worst-case SSNR error. In the first set of tests, the SSNR was set to 1 dB and the telemetry configuration tests are shown in Table 1. It includes eight tests ranging in symbol rate from 32 symbols per sec (sps) typical of Pioneer 10 to 537.6 kbps, which corresponds to the high data channel of Magellan. In each test,

the SERSNR was deduced from the measured SER and the loss was computed based on the assumption that the input SSNR is exactly 1 dB. The corresponding SERSNR loss is computed in each case and depicted in Fig. 3(a). Simultaneous measurements of the BER were made by the SPT and the corresponding BERSNR was computed using the performance of the (7,1/2) convolutional code, as given in Fig. 2(b). The resulting BERSNR loss is shown in Fig. 3(b) for all tests. Note that the SERSNR loss for all three systems agrees to within 0.2 dB for all symbol rates between 2.4 kbps and 268.8 kbps. For 32 sps and 80 sps (or the equivalent 16 bps and 40 bps), the ARX II outperformed the BLK-III/BPA and BLK-III/SDA/SSA combinations by as much as 1.2 dB in SSNR. The reason for this improvement is the capability of the ARX II to use narrower carrier loop bandwidths than the Block-III Receiver, resulting in a higher loop SNR and smaller phase jitter.

At the highest rate of 537.6 kbps, the ARX-II performance was degraded by about 0.4 dB with respect to the BLK-III/BPA combination (the SSA is specified to operate at a maximum of 200 kbps and was not operational in this case). Other tests were performed at a lower SSNR (–1 dB), and the results are shown in Fig. 4 with the test configuration in Table 2. The subcarrier and symbol synchronization loops were operating with the same bandwidths as those given in Table 1. Here again, the ARX-II SERSNR loss, Fig. 4(a), and the BERSNR loss, Fig. 4(b), at 537.6 kbps are at least 0.4 dB worse than those of the BLK-III/BPA. Furthermore, the two figures are reflecting the coding gain of 3 dB to within 0.2 dB.

The telemetry degradation of the ARX II at 537.6 kbps was of primary concern and has been thoroughly investigated. It turned out that the additional loss was really due to the filtering operation on the signal in the exciter itself, and not in the ARX II. This is the reason why the loss of the BLK-III/BPA combination also increases at that symbol rate. Figure 5 illustrates the filtering phenomenon clearly. In Fig. 5(a), the spectrum of a 1-MHz square-wave subcarrier is shown indicating only the presence of the odd harmonics. The bandwidths of the various signals and systems are indicated in Fig. 5(b); the exciter has a 2-MHz bandwidth which passes only the first harmonic of the 960-kHz subcarrier. The Block-III Receiver and the ARX II have roughly 4-MHz and 8-MHz telemetry bandwidths, respectively. Expanding a square-wave subcarrier in a Fourier series, one obtains

$$\sin(2\pi f_{sc}t) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \sin[2\pi(2n+1)f_{sc}t] \quad (2)$$

where  $f_{sc}$  is the subcarrier frequency. The power of any odd harmonic is proportional to  $1/n^2$ , as expected. The exciter limits both the BLK-III/BPA and the ARX II to the first harmonic of the signal. In the Block-III Receiver, the noise bandwidth allows the first and third harmonics of the reference subcarrier to heterodyne the noise to baseband, resulting in a noise degradation of

$$\text{BLK-III noise degradation} = \frac{1}{1 + 1/3^2} = 0.9 = -0.46 \text{ dB} \quad (3)$$

On the other hand, the ARX-II noise bandwidth allows the first through the ninth harmonics of the reference subcarrier to heterodyne the noise to baseband, resulting in a noise degradation of

$$\begin{aligned} \text{ARX-II noise degrad.} &= \frac{1}{1 + 1/3^2 + 1/5^2 + 1/7^2 + 1/9^2} \\ &= 0.845 = -0.73 \text{ dB} \end{aligned} \quad (4)$$

Removing those degradations from the measured SERSNR loss for both the BLK-III/BPA and ARX II, one obtains the dashed lines in Fig. 3(a) at 537.6 kbps, which are in total agreement with the SERSNR loss at lower data rates. If the exciter bandwidth limitations had been known before the tests, the received subcarrier in the ARX II would have been mixed with a pure sine wave, resulting in 0.73-dB improvement over the measured performance of the ARX II with a square wave and a 0.46-dB improvement over the BLK-III/BPA combination. The ARX II can heterodyne the received subcarrier with a square wave, a sine wave, or a filtered square wave such that only the first  $K$  harmonics are passed. As an example, the received subcarrier can be mixed with the combination of first and third harmonics only. All this processing can be controlled by a software command and can be changed on the fly. In case the filtering on the subcarrier is unknown to the receiver, several combinations can be experimented with in the receiver in real time and the one providing the best performance utilized. This additional flexibility is easily obtained in digital implementations and guarantees that the local subcarrier is “matched” to the received one.

### C. Performance of the SNR Estimators

Recall that one of the objectives of these tests was to assess the performance of the ARX-II SSNR estimator. During these tests, the SNR estimator of the ARX II suffered from a slight hardware problem that produced a bias

in the SNR estimates. The problem resulted from two missing samples in the signal power estimate, thus, degrading more at the higher symbol rates. The problem was not fixed in time for these tests, but will be fixed in the next board version. Figure 6 depicts the SSNR loss as estimated by the SSNR estimator and the SERSNR estimator for both the ARX II and the BLK-III/BPA combination. Both estimators seem to agree to within 0.2 dB in the BLK-III/BPA combination, except at the lowest rate where the difference was about 0.6 dB. In the ARX-II case, the agreement was within roughly 0.4 dB, except at the highest rate where the estimates diverged by as much as 1 dB. More tests will be conducted in the next implementation of the estimator to obtain better results.

### D. Performance of the ARX II With Pioneer 10's Signal

The last set of tests were specific to the Pioneer 10 spacecraft, which transmits the lowest symbol rate of 32 sps with the weakest received signal of roughly  $P_T/N_0 = 20$  dB-Hz using the 70-m antennas. The telemetry configuration tests for Pioneer 10 are shown in Table 3, where three different values of  $P_C/N_0$  were used, namely, 13 dB-Hz, 11 dB-Hz, and 9 dB-Hz. The ability of narrowing the carrier loop bandwidth of the ARX II is clearly shown in Fig. 7(a), where as much as an additional 1.4-dB improvement in SSNR has been obtained over the BLK-III/BPA combination. In CTA21, the optimum loop bandwidth obtained was about 0.5 Hz due to the possible presence of significant phase noise. Some of these measurements were repeated in the laboratory and the optimum loop bandwidth was further narrowed to about 0.25 Hz, adding credence to the theory of significant phase noise present in CTA21. Figure 7(b) depicts the measured SERSNR loss and the theoretical losses assuming thermal noise only. The carrier loss was computed from Fig. 8<sup>1</sup> for the different loop SNRs. For the subcarrier and symbol synchronization, SSNR degradation was computed,<sup>2</sup> respectively,

$$\begin{aligned} D_{sc} &= 1 - 1.0159\sigma_{sc}, \quad \sigma_{sc}^2 = 1/\rho_{sc}, \\ \rho_{sc} &= \frac{P_D}{N_0 B_{sc}} \left(\frac{2}{\pi}\right)^2 \frac{1}{1 + N_0/2E_s} \end{aligned} \quad (5)$$

<sup>1</sup> *Deep Space Network/Flight Project Interface Design Handbook, Vol. I: Existing DSN Capabilities*, JPL 810-5, Rev. D (internal document), Jet Propulsion Laboratory, Pasadena, California.

<sup>2</sup> J. Statman, *DSN Receiver Losses in Galileo Experiment*, JPL Interoffice Memorandum 331-88.5-050 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 14, 1988.

$$D_{sy} = 1 - 0.2534\sigma_{sy}, \quad \sigma_{sy}^2 = 1/\rho_{sy},$$

$$\rho_{sy} = \frac{P_D}{N_0 B_{sy}} \frac{\operatorname{erf}^2(\sqrt{E_s/N_0})}{2\pi^2} \quad (6)$$

These formulas assume that no windowing is used in the loops and that the windowing needs to be accounted for by multiplying the loop variance by its corresponding relative window size  $W$ . The symbol synchronization loop was operated with  $W_{sy} = 1/2$  and the window is accounted for in Table 4, which compares the various synchronization effects for various operating points in the test table. The subcarrier and symbol degradations are significant and surpass the carrier loss in some instances. These degradations can be reduced either by narrowing the loop bandwidths or by employing windows [4].

During the course of these tests, the ARX-II subcarrier and symbol lock detectors were not tested extensively and, as a result, require further testing. The scaling of soft symbols from the ARX II to the TPA was performed manually and needs to be automated in the next upgrade. To test the performance of future systems with a 0.1-dB accuracy, CTA21 test equipment needs to be upgraded in order to provide an accuracy better than 0.1 dB (currently, it is 0.5 dB). Furthermore, the CTA21 test capability does

not support SER testing in excess of 1 Msps (537.6 kbps required recording and playback of data). All of these issues are areas of concern that need to be addressed in the future.

## V. Goldstone Tests

Telemetry tests with the ARX II were performed at Goldstone in early November 1990 by tracking and demodulating signals from the International Comet Explorer (ICE), Voyager 1, Magellan low and high data rate channels, and Pioneers 10 and 11. Table 5 depicts the signal characteristics of each spacecraft tracked. Early results seem to agree with those run in CTA21, but more extensive analysis of the data is planned.

## VI. Conclusion

The CTA21 tests have demonstrated the telemetry capability of the ARX II to demodulate signals with symbol rates up to 537.6 kbps. Test results specific to Pioneer 10 show a 1.2- to 1.4-dB SERSNR improvement in the telemetry performance of the ARX II versus the BLK-III/BPA combination due to the capability of the ARX II to use narrower loop bandwidths.

## Acknowledgments

The authors thank Dr. J. K. Holmes and Mr. J. Berner for their valuable advice in analyzing the data. Also, the help of both Mr. C. Pasqualino and Mr. R. Labelle is greatly appreciated in setting up the ARX II. In addition, the support of the personnel in CTA21, especially Mr. C. Brown, is acknowledged.

## References

- [1] S. Hinedi, "A Functional Description of the Advanced Receiver," *TDA Progress Report 42-100*, vol. October-December 1989, pp. 131-149, Jet Propulsion Laboratory, Pasadena, California, February 1990.
- [2] S. Hinedi, R. Bevan, H. Del Castillo, P. Kinman, D. Chong, and R. Labelle, "Digital Doppler Extraction Demonstration with the Advanced Receiver," *TDA Progress Report 42-100*, vol. October-December 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 160-173, February 1990.

- [3] M. K. Simon and A. Mileant, "SNR Estimation for the Baseband Assembly," *TDA Progress Report 42-85*, vol. January–March 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 118–126, January 1986.
- [4] W. Hurd and S. Aguirre, "A Method to Dramatically Improve Subcarrier Tracking," *TDA Progress Report 42-86*, vol. April–June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 103–110, April 1986.

Table 1. Telemetry configuration test table (SSNR = 1 dB and BSNR = 4 dB)

Test number	Bit rate, bps	Symbol rate, sps	$\Delta$ Mod. index, deg	Subcarrier frequency, kHz	Input SSNR, dB	$P_D/N_0$ , dB-Hz	$P_C/N_0$ , dB-Hz	Carrier loop bandwidth ( $2 B_C$ )			Subcarrier bandwidth, $B_{SC}$ , Hz			Symbol loop bandwidth, $B_{SY}$ , Hz		
								ARX II	BLK III	ARX II	BPA	SDA	ARX II	BPA	SDA	ARX II
1	16	32	65.9	32.768	1	16.0	9.0	1	0.1	0.04	0.06	0.1	0.04	0.1	0.04	0.3
2	40	80	45.0	22.5	1	20.0	20.0	2	0.2	0.07	0.08	0.1	0.07	0.1	0.07	0.4
3	1,200	2,400	70.0	360.0	1	34.8	26.0	6	0.2	0.15	0.12	0.2	0.15	0.2	0.15	0.1
4	21,600	43,200	80.0	360.0	1	47.4	32.3	6	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.5
5	43,200	86,400	80.0	360.0	1	50.4	35.3	6	0.2	0.2	0.2	0.2	0.2	0.2	0.2	1.0
6	67,200	134,400	80.0	360.0	1	52.2	37.2	6	0.2	0.2	0.2	0.2	0.2	0.2	0.2	1.5
7	134,400	268,800	80.0	360.0	1	55.2	40.2	6	0.2	0.2	0.2	0.2	0.2	0.2	0.2	N/A
8	268,800	537,600	80.0	960.0	1	58.2	43.2	6	0.2	0.2	0.2	0.2	0.2	0.2	0.2	N/A

Table 2. Telemetry configuration test table (SSNR = -1 dB and BSNR = 2 dB)

Test number	Bit rate, bps	Symbol rate, sps	Mod. index, deg	Subcarrier frequency, kHz	Input SSNR, dB	$P_D/N_0$ , dB-Hz	$P_C/N_0$ , dB-Hz	Carrier loop bandwidth ( $2 BL$ )		
								ARX II	BLK III	BLK III
1	1,200	2,400	70	360	-1	32.8	24.0	6	6	13
2	21,600	43,200	80	360	-1	45.4	30.3	6	6	20
3	43,200	86,400	80	360	-1	48.4	33.3	6	6	22
4	67,200	134,400	80	360	-1	50.2	35.2	6	6	24
5	134,400	268,800	80	360	-1	53.2	38.2	6	6	24
6	268,800	537,600	80	960	-1	56.2	41.2	6	6	24

Table 3. Telemetry configuration test table for Pioneer 10

Test number	Symbol rate, sps	$\Delta$ Mod. index, deg	Subcarrier frequency, kHz	Input SSNR, dB	$P_D/N_0$ , dB-Hz	$P_C/N_0$ , dB-Hz	BLK-III carrier margin, dB	Carrier loop bandwidth (2 BL)		ARX-II carrier loop SNR, dB	Subcarrier loop bandwidth, Hz	Symbol loop bandwidth, Hz
								BLK III	ARX II			
1	32	65.9	32.768	5	20	13	9	4.4	4.4	9.6	0.1	0.1
2								3.0		11.2		
3								1.0		16.0		
4								0.5		19.0		
5	32	65.9	32.768	3	18	11	7	4.0	4.0	8.0	0.1	0.1
6								3.6		8.5		
7								3.0		9.2		
8								1.0		14.0		
9								0.5		17.0		
10	32	65.9	32.768	1	16	9	5	3.3	3.0	7.2	0.1	0.1
11								1.0		12.0		
12								0.5		15.0		

**Table 4. Various synchronization losses for Pioneer 10**

Configuration parameters, dB	Symbol degradation, dB	Subcarrier degradation, dB	Carrier loop SNR, dB	Carrier loss, dB	Telemetry loss, dB
$E_s/N_0 = 5$	0.12	0.26	9.6	1.00	1.38
			11.2	0.45	0.83
$P_C/N_0 = 13$	0.12	0.26	16.0	0.07	0.45
			19.0	0.05	0.43
$E_s/N_0 = 3$	0.15	0.33	14	0.3	0.78
$P_C/N_0 = 11$					
$E_s/N_0 = 1$	0.21	0.44	0.44	1	1.65
$P_C/N_0 = 9$					

**Table 5. Goldstone tests**

Spacecraft	DSS	RF frequency	$P_C/N_0$ , dB-Hz	Subcarrier frequency, kHz	Bit rate, bps	Symbol rate, sps	SSNR, dB
Ice	14	S-band	21	1.024	128	256	1
Voyager 1	15	X-band	25	360	600	1200	2
Magellan	15	X-band	48	960	268,800	537,600	3
				22.5	40	80	14

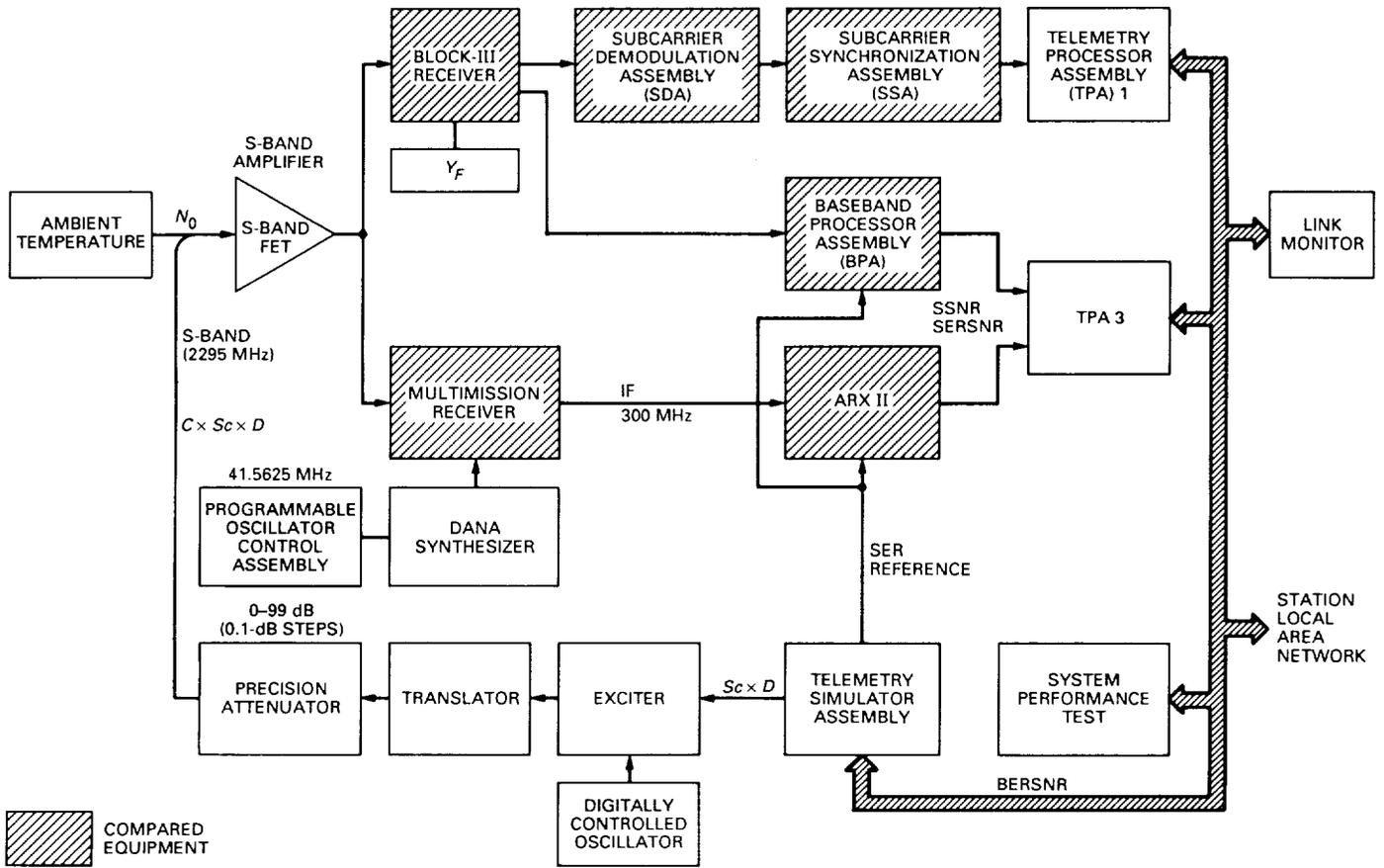


Fig. 1. ARX-II telemetry configuration test at CTA21.

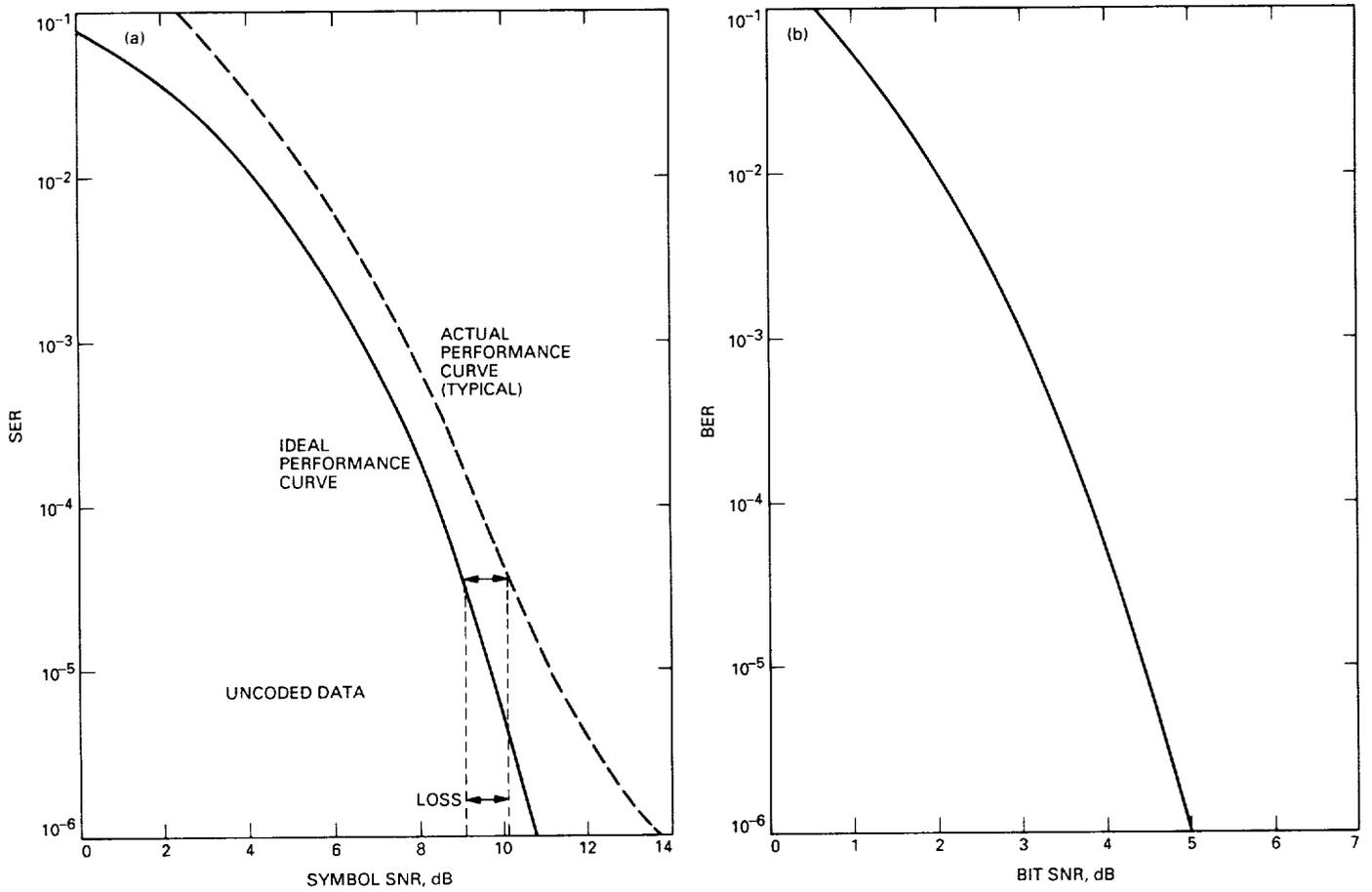


Fig. 2. Signal-to-noise ratio performance for: (a) telemetry loss definition, and (b) the (7,1/2) convolutional code using Viterbi decoding ( $Q = 3$  bits).

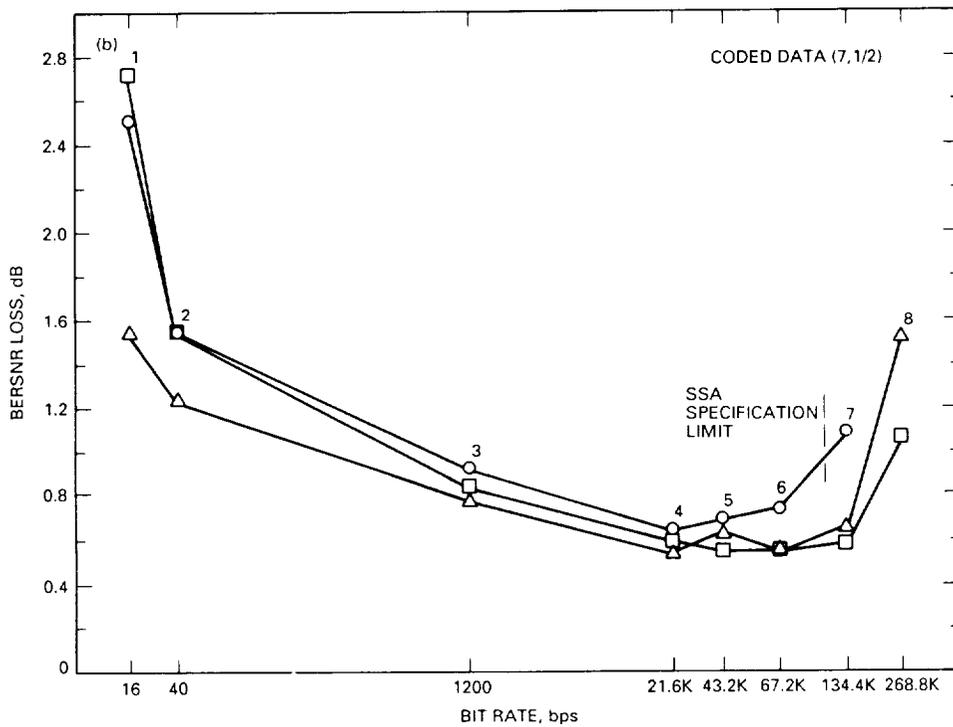
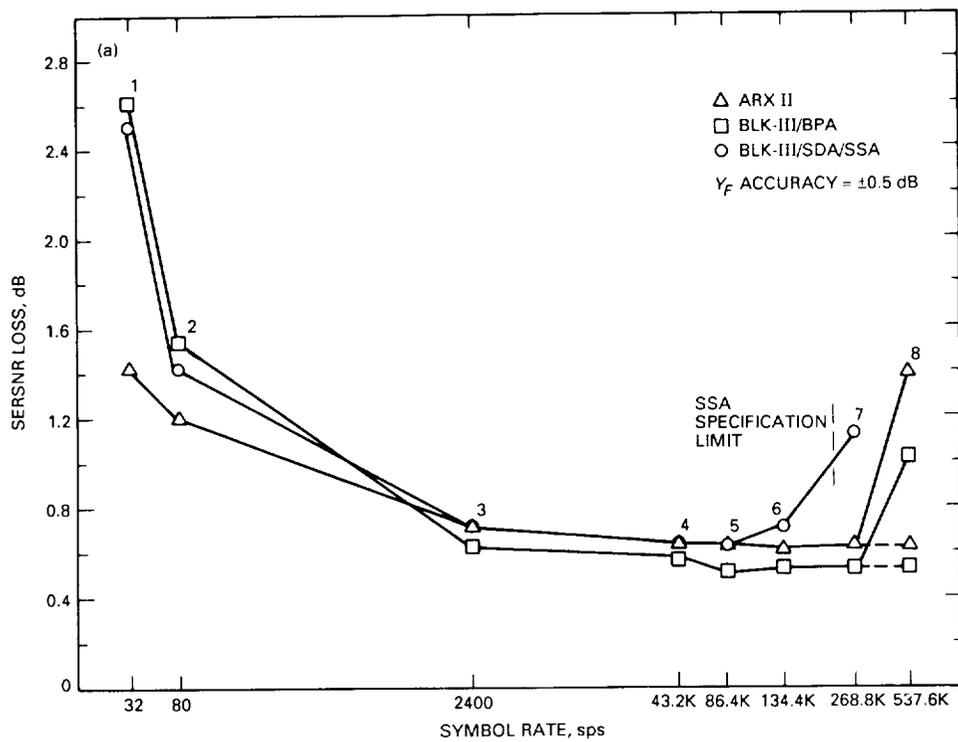


Fig. 3. Signal-to-noise ratio loss for: (a) SER versus symbol rate (SSNR = 1 dB), and (b) BER versus bit rate (BSNR = 4 dB).

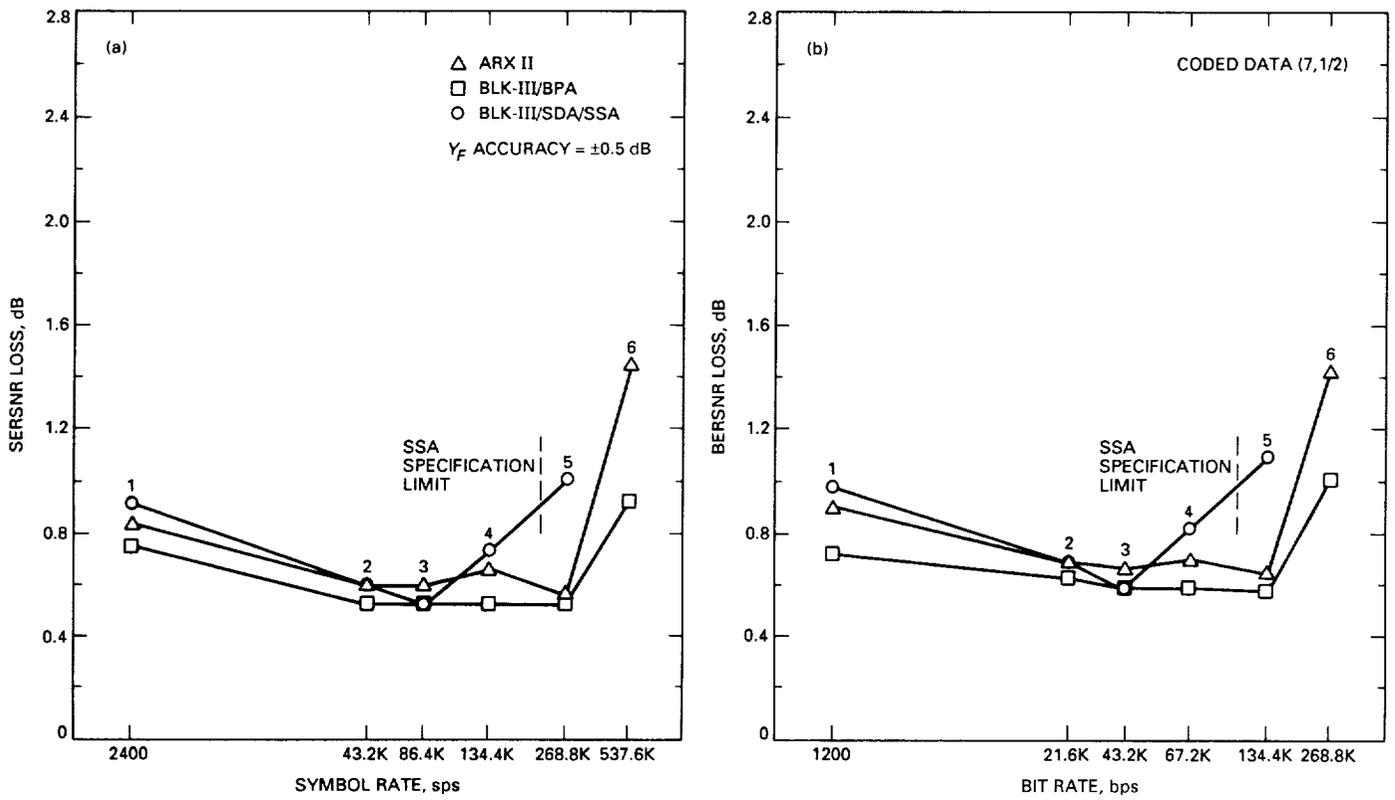


Fig. 4. Signal-to-noise ratio loss for: (a) SER versus symbol rate (SSNR = -1 dB), and (b) BER versus bit rate (BSNR = 2 dB).

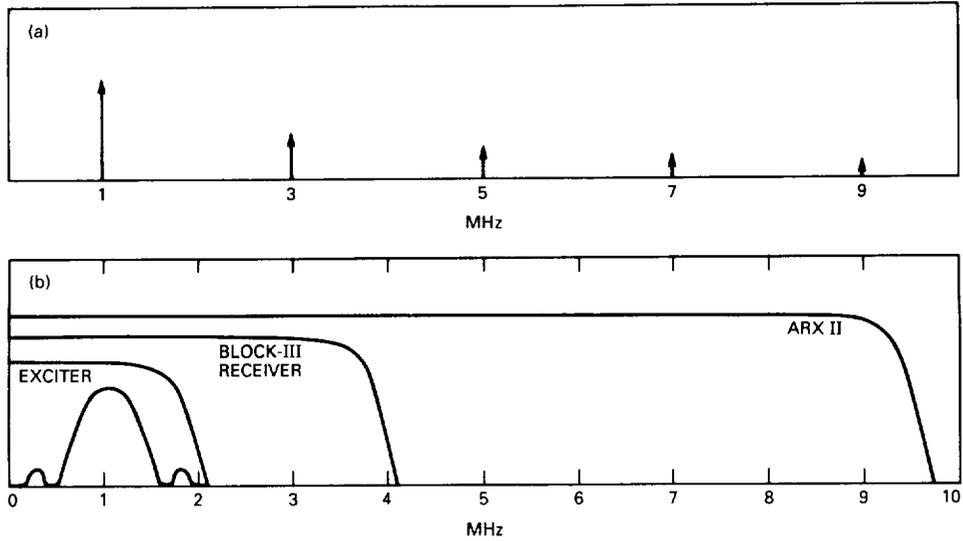


Fig. 5. Comparison of the spectrums of: (a) a 1-MHz square-wave subcarrier, and (b) the exciter, the Block-III Receiver, and the ARX II.

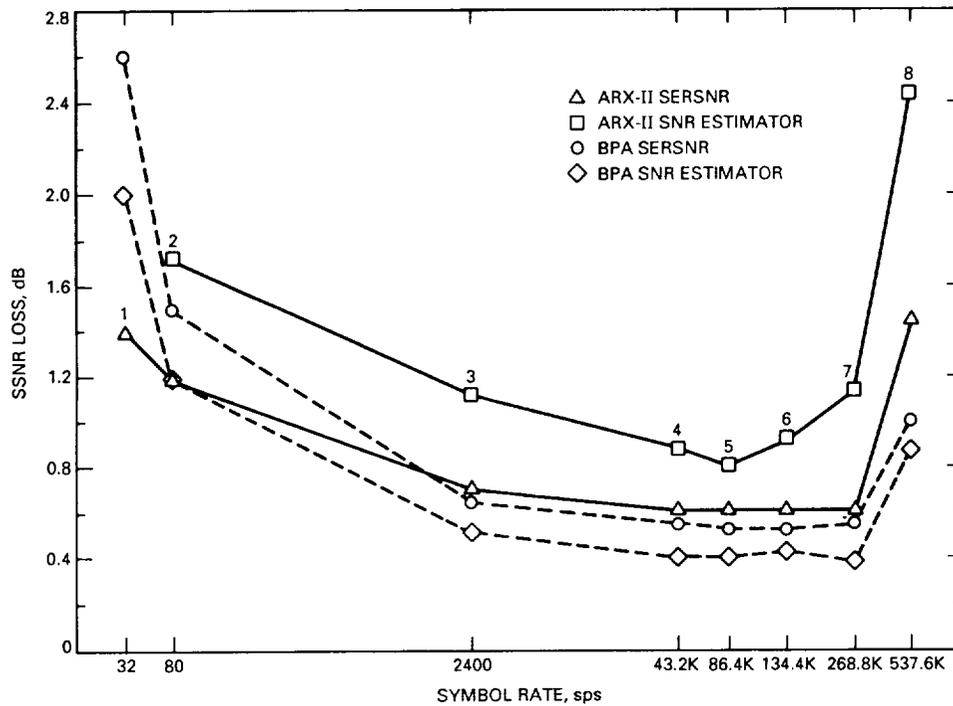


Fig. 6. SSNR and SERSNR losses versus symbol rate (SSNR = 1 dB).

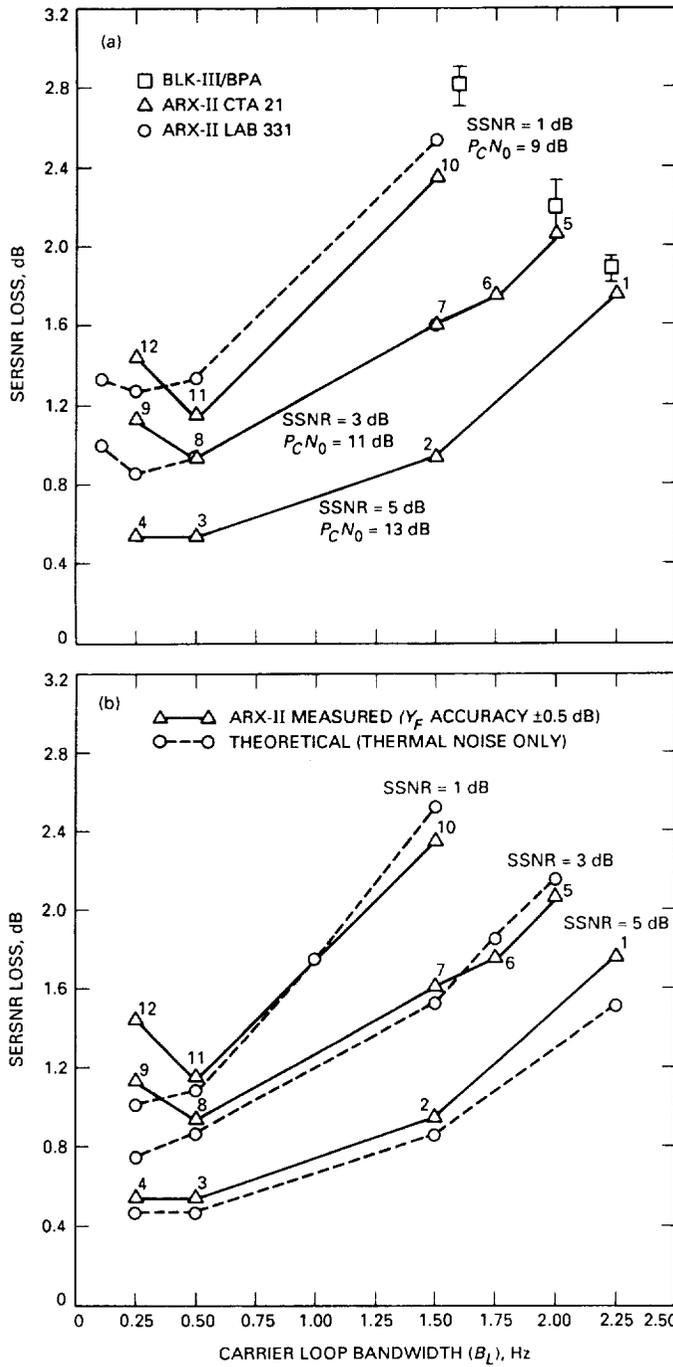


Fig. 7. Comparison of SERSNR losses: (a) versus carrier loop bandwidth, and (b) measured versus theoretical losses for Pioneer 10.

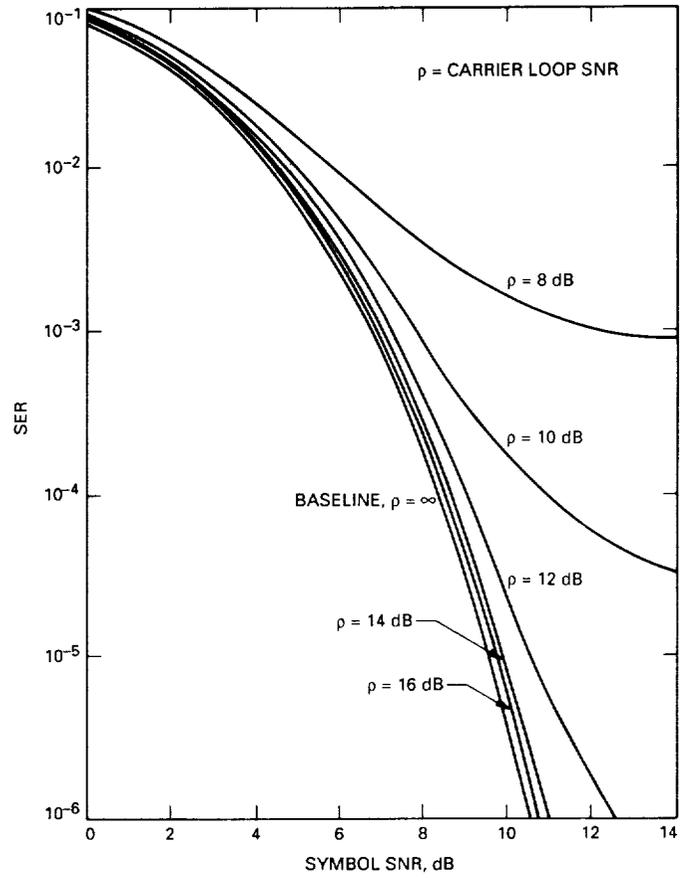


Fig. 8. SER versus symbol SNR in the presence of carrier tracking phase error.

## Appendix A

### Modulation Index Setup Procedure

This procedure is used to adjust the modulation index in the telemetry tests. Let  $P_T$  denote total received signal power. Then the carrier and data powers are given respectively by

$$P_C = P_T \cos^2 \Delta \quad (\text{A-1})$$

$$P_D = P_T \sin^2 \Delta \quad (\text{A-2})$$

where  $\Delta$  denotes the modulation index in radians. The Block-III Receiver contains an automatic gain control (AGC) loop which maintains a nearly constant carrier power at the receiver's output. The carrier power is measured in a narrow bandwidth, typically a few hertz. The voltage which controls the receiver gain is measured by

a digital voltmeter. The modulation index is set up as follows:

- (1) The transmitter is turned on with a "very strong" carrier (all modulation off).
- (2) The receiver AGC voltage is read (AGC1).
- (3) A precision attenuator is added equal to the value of the carrier suppression, which is  $20 \log_{10}(\cos \Delta)$ .
- (4) The receiver AGC voltage is read (AGC2).
- (5) The precision attenuator is adjusted until the AGC voltage is that of step (2), i.e., AGC1.
- (6) The modulation is turned on.
- (7) The modulation amplitude is adjusted at the TSA until the AGC voltage reads the value of step (4), i.e., AGC2.

## Appendix B

### Y-Factor Setup Procedure

The Y-factor setup procedure is used to adjust the SSNR of the received signal. It is used at both the Telecommunications Development Laboratory (TDL) and in CTA21. The procedure is very straightforward and relies on measuring the total power (signal plus noise) of a pure tone in a known bandwidth and comparing it to the noise power alone in the same bandwidth. The Y-factor (in decibels) is defined as

$$Y_F = 10 \log_{10}(1 + 10^{\frac{SSNR}{10}}) \quad (\text{B-1})$$

where (also in decibels)

$$SSNR = (SNR + 10 \log_{10} R) - (20 \log_{10} \sin \Delta + 10 \log_{10} B) \quad (\text{B-2})$$

where  $R$  is the symbol rate and  $B$  the bandwidth of the Y-factor detector and  $A$  is an attenuation pad. Small values of SSNR can be precisely calibrated by Y-factoring an artificial SSNR which is much higher than the desired SSNR. An “add-pad” is then introduced between the transmitter and the receiver after the procedure is completed. Since the “add-pad” is a discrete-step attenuator, its finite selectable attenuation settings can be accurately calibrated beforehand and it introduces insignificant error to the Y-factor technique.

Intuitively, the Y-factor is the ratio of signal power plus noise power to the noise power in some bandwidth. It can be expressed as  $(N_0B + P_T)/N_0B$  where  $P_T$  is the total signal power (typically the carrier power as the Y-factor procedure is performed with tones and no modulation). The equipment configuration used for Y-factoring is shown in Fig. 1. One of the outputs of the Block-III Receiver is

a 50-MHz IF signal with no AGC, which is used in the Y-factor procedure. The steps in the calibration are:

- (1) Calculate the Y-factor for the desired SSNR from Eq. (B-1).
- (2) Turn the transmitter off.
- (3) With the Y-factor attenuator at some reference, measure the power level with a power meter, i.e., measure the noise power in a known bandwidth  $B$ .
- (4) Increase the Y-factor attenuator by the value of the Y-factor computed in step (1).
- (5) Turn the transmitter on with all modulation off.
- (6) Adjust the the RF precision attenuator until the power meter of the Y-factor reads the reference level of step (3).
- (7) Add an attenuation pad in the precision attenuator, if applicable.
- (8) Turn the modulation on.

Basically, the Y-factor procedure measures first noise power ( $N_0B$ ), in the absence of signal, with a power meter. An attenuation exactly equal to the Y-factor is added to decrease the power measurement accordingly. When a tone is injected, the power meter (which is measuring signal plus noise) should read exactly  $N_0B$  if the signal power is at the right level, i.e., the power meter is reading  $N_0B + P_T$  or  $N_0B(1 + P_T/N_0B)$  or  $N_0B Y_F$ , but because of the attenuation by  $Y_F$ , the reading becomes  $N_0B$ . If the signal is not at the right level, the signal level is adjusted with the RF precision attenuator until the power meter reads  $N_0B$ . Actually, the Y-factor procedure produces the correct  $P_C/N_0$  and, hence, the correct SSNR, since the modulation index was set as in Appendix A.

5/3-32  
43271  
N91-18321  
F-4  
JJ574039

# On Estimating the Phase of a Periodic Waveform in Additive Gaussian Noise— Part III

L. L. Rauch

Office of Telecommunications and Data Acquisition

*Motivated by advances in signal processing technology that support more complex algorithms, researchers have taken a new look at the problem of estimating the phase and other parameters of a nearly periodic waveform in additive Gaussian noise, based on observation during a given time interval. In Part I, the general problem was introduced and the maximum a posteriori probability criterion with signal space interpretation was used to obtain the structures of optimum and some suboptimum phase estimators for known constant frequency and unknown phase with an a priori distribution. In Part II, optimal algorithms were obtained for some cases where the phase (and frequency) is a parameterized function of time with the unknown parameters having a joint a priori distribution. The intrinsic and extrinsic geometry of hypersurfaces was introduced to provide insight to the estimation problem for the small-noise and large-noise cases. In Part III, the actual performances of some of the highly nonlinear estimation algorithms of Parts I and II are evaluated by numerical simulation using Monte Carlo techniques.*

## I. Introduction

The work of Part I [1] and Part II [2] is limited to analytical results which, although they provide the structure of nonlinear optimum estimators, can give the performance of these estimators only in the case of small noise—and then not always in the case of some suboptimum estimators such as phase-locked loops. In what follows, certain nonlinear estimators are exactly simulated by numerical methods and their performance is evaluated

by Monte Carlo techniques. These include nonsinusoidal waveforms with unknown phase and known frequency and sinusoidal waveforms with unknown phase and frequency.

## II. Performance of Phase Estimators With Known Constant Frequency

For a sinusoid of known frequency and unknown phase, the probability density function given in footnote 2 on

page 157 of [1] is easily numerically integrated, after multiplication by error squared, to obtain the mean-square (ms) phase error of the well-known optimum estimator of Eq. (33) of [1]. For a nonsinusoid consisting of a fundamental and one or more harmonics, the optimum estimator is the maximization of Eq. (61) of [1]. Since the probability distribution of the error is not available, the ms error of this estimator is found by numerical methods using a Monte Carlo technique. The results for (a) a sinusoid, (b) a sinusoid plus third square-wave harmonic, and (c) a sinusoid plus third, fifth, and seventh square-wave harmonics are plotted in Fig. 1. As the maximum slope of the periodic waveform increases with additional harmonics, the large-noise threshold of the optimum estimator becomes steeper and moves to higher input signal-to-noise ratios, as would be expected. Some insight is given by Fig. 2 of [2] and Sections VI of [1] and VIII of [2].

It is interesting to compare the performance of a second-order phase-locked loop (PLL) with that of the previous optimum estimator for the phase of a sinusoid with known frequency. The PLL filter, preceding the voltage-controlled oscillator, is taken to be of the form  $(\tau_1 s + \tau_2)/s$ , where  $\tau_1$  and  $\tau_2$  are chosen to give a damping ratio of  $1/\sqrt{2}$ . This is essentially what is used in the carrier tracking loops of Deep Space Network (DSN) receivers.<sup>1</sup> The noise bandwidth of the optimum estimator is  $1/2T$ , where  $T$  is the duration of the observation interval.

For a valid comparison between the PLL and the optimum estimator, the PLL must be observed at a time interval  $T$  after the PLL is turned on, with an initial phase error uniformly distributed over one cycle (and no initial frequency error). If the PLL filter is chosen to give a PLL noise bandwidth equal to that of the optimum estimator, this turns out to be too small. In this case, the ms phase error of the PLL is dominated by the transient responses of the loop to the initial phase errors. This is much larger than the ms phase error contribution resulting from the additive noise for any useful signal-to-noise ratio. The observation interval  $T$  is only  $\sqrt{2}/3\tau \approx 0.150$  of the undamped period of the PLL when its noise bandwidth is equal to that of the optimum estimator ( $1/2T$ ). As the loop noise bandwidth is increased (period is decreased) the ms phase error contribution from the initial transients

<sup>1</sup> Actually, the denominator of the DSN PLL filter is of the form  $s + \tau$ , where  $\tau$  is on the order of 0.001 at the smaller loop bandwidths.

decreases while the contribution from the additive noise increases. For each input signal-to-noise ratio there is an optimum loop noise bandwidth which minimizes the total ms phase error of the PLL at the end of the observation interval  $T$ .

In Fig. 2 these PLL minimum ms phase error values are plotted together with the ms phase error of the optimum estimator. Even at large input signal-to-noise ratios (small noise) the PLL performance is about 9 dB worse than the optimum estimator. The PLL results are obtained by numerical solution (fourth-order Runge-Kutta) of the base-band second-order nonlinear differential equation for the loop phase error.

### III. Performance of Optimum Estimator With Unknown Constant Frequency

In this case the estimation algorithm consists of choosing the frequency  $f$  in Eq. (102) of [2] to maximize Eq. (113), supported by Eq. (103). The phase is then given by Eq. (111). For this numerical simulation, the a priori distribution of phase is uniform over one cycle and the independent a priori distribution of frequency is taken to be uniform over the interval  $(-5.5/T$  to  $+5.5/T)$  centered around a given frequency. [The index in Eq. (103) runs from  $-5$  to  $+5$ .]

The ms phase error of the optimum estimator is given in Fig. 3, with the ms phase error for known frequency as reference. In accordance with the analytical result of Eq. (96) of [2], the small-noise performance is the same as that for known frequency. However, the steeper large-noise threshold occurs at a higher signal-to-noise ratio.

The ms frequency error of the optimum estimator is given in Fig. 4. It is evident that the large-noise threshold is more abrupt than that for phase error. In the case of unknown phase and frequency, the PLL performance is much worse than for the case of unknown phase and known frequency, shown in the previous section.

It should be kept in mind that the results of this section depend on the a priori distribution, above, chosen for the unknown frequency.

## References

- [1] Rauch, L. L., "On Estimating the Phase of a Periodic Waveform in Additive Gaussian Noise—Part I," *DSN Progress Report 42-45*, vol. March and April 1978, Jet Propulsion Laboratory, Pasadena, California, pp. 152-164, June 15, 1978.
- [2] Rauch, L. L., "On Estimating the Phase of a Periodic Waveform in Additive Gaussian Noise—Part II," *TDA Progress Report 42-79*, vol. July-September 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 17-24, November 15, 1984.

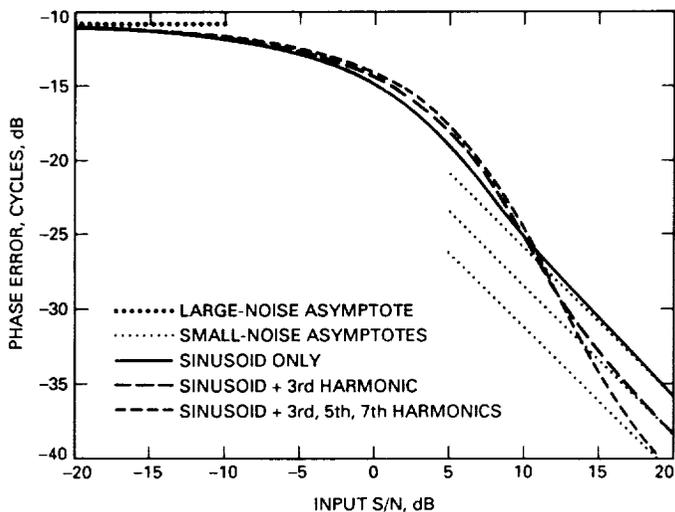


Fig. 1. Phase-error performance of three optimum estimation algorithms for the case of known constant frequency. The large-noise asymptote corresponds to phase error uniformly distributed over one cycle. The three small-noise asymptotes come from Eqs. (54), (59), and a similar equation for the first four components of a square wave, obtained by using Eq. (56) of [1].

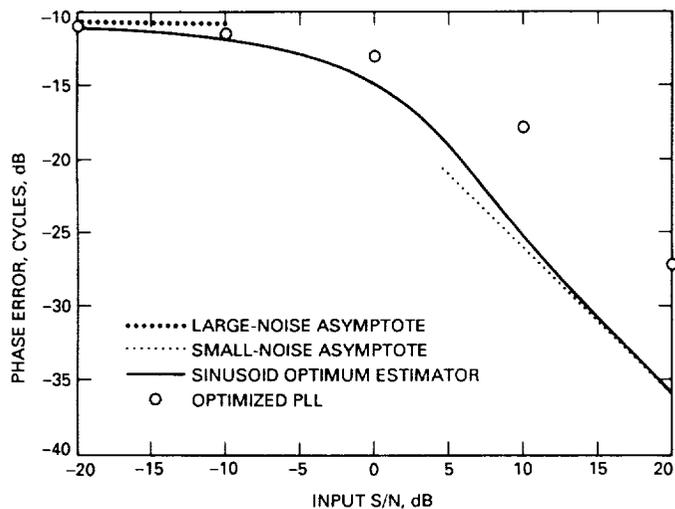


Fig. 2. Phase-error performance of an optimized phase-locked loop subject to the same signal-to-noise ratio and observation interval as the optimum estimator.

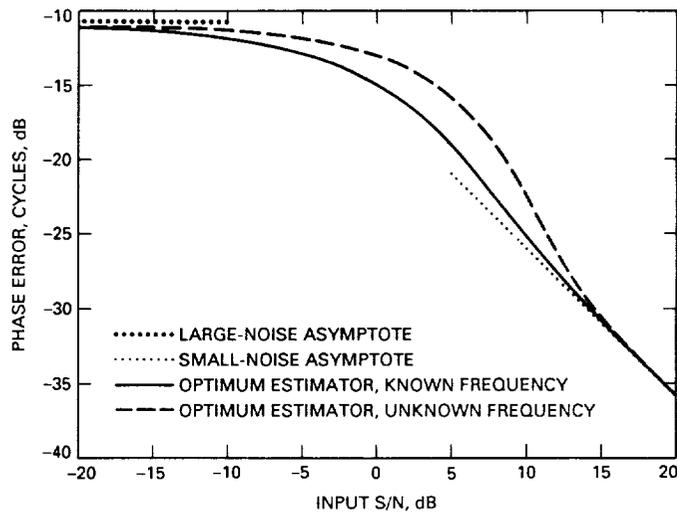


Fig. 3. Phase-error performance of the optimum estimator for the case of unknown constant frequency. The small-noise asymptote comes from Eq. (96) of [2].

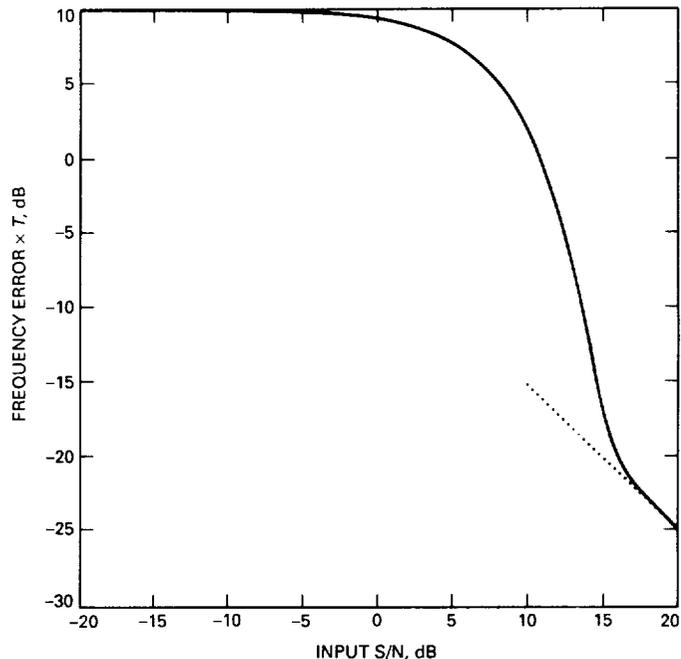


Fig. 4. Frequency error performance of the optimum estimator for the case of unknown constant frequency. The small-noise asymptote comes from Eq. (99) of [2]. The large-noise asymptote (not plotted) is at 10.04 dB. This corresponds to a frequency error uniformly distributed from  $-5.5/T$  to  $+5.5/T$ .

514-32  
332 1/2  
N91-183224

JJ 574450

# Digital Test Signal Generation: An Accurate SNR Calibration Approach for the DSN

B. O. Gutierrez-Luaces  
Telecommunications Systems Section

*A new method of generating analog test signals with accurate signal-to-noise ratios (SNRs) is described. High accuracy will be obtained by simultaneous generation of digital noise and signal spectra at a given baseband or bandpass-limited bandwidth. The digital synthesis will provide a test signal embedded in noise with the statistical properties of a stationary random process. Accuracy will only be dependent on test integration time with a limit imposed by the system quantization noise (expected to be 0.02 dB). Setability will be approximately 0.1 dB. The first digital SNR generator to provide baseband test signals is being built and will be available in early 1991.*

## I. Introduction

Spacecraft link performance optimization has always been a requirement for the missions supported by the Deep Space Network (DSN). This optimization relies on accurate predictions for the degradations (and losses) encountered in the different modulation and detection processes in use in the telecommunications link. Mathematical models for these processes are usually available beforehand, to be later verified by tests run at the Telecommunications Development Laboratory (TDL), the Compatibility Test Area, and sometimes at the DSN stations.

Figure 1 represents the Signal Processing Center (SPC) portion of a typical DSN communications link. The parameters of interest to be considered are:

$$\text{Carrier SNR} = P_C/N_0 \quad (1)$$

$$\text{Symbol SNR} = ST_S/N_0 \quad (2)$$

$$\text{Bit SNR} = ST_b/N_0 \quad (3)$$

with

$P_C$  = the portion of received power in the residual carrier

$S$  = the portion of received power in the modulation sidebands

$T_S$  = the symbol period

$T_b$  = the bit period

$N_0$  = the noise spectral density

Note that the above parameters, to be referred to generically as SNR, may also be considered as the residual

carrier power, energy per symbol, and energy-per-bit to spectral-noise-density ratios, respectively.

Calibration of telemetry system losses and SNR degradation<sup>1</sup> has traditionally been performed with the well-known Y-factor method [1]. Different error sources come into play depending on whether the calibration is performed in the carrier or in the modulated part of the transmitted spectrum. In [1], these error sources have been identified and quantified with a final worst-case error of  $\pm 0.45$  dB. Reported accuracy of the manual method varies from 0.3 dB at the TDL to 1 dB at the DSN stations.

To improve the accuracy and to automate the measurement process, a method has been developed to digitally synthesize the precision test signal. The signal-to-noise ratios obtained are then independent of gain variations, and full knowledge and control of waveshapes and the modulation index can be achieved. As shown in Fig. 2, calibration of the degradation and losses of the different processes are performed as in the traditional Y-factor method. Calibrated SNR and bit-error rate (BER) of signals input to the device under test are compared to the observed SNR and BER of the signals the device outputs. The advantage of the digital SNR generator (DSG) approach is that the SNR-generation process is truly stationary. Therefore, long integration times can be used to obtain the accuracy and precision needed. Details of this approach will be described in the following paragraphs.

## II. DSG Description

The DSG, as shown in Fig. 2, will generate test signals at baseband (BB). The design allows the generation of an intermediate frequency (IF) to be upconverted, as a future option, to the desired radio frequency band. Note that the baseband spectrum may correspond to coded or uncoded data. Figure 3 is a flowchart describing programmable capabilities of the DSG. It also shows the basic software functions for the baseband spectrum generation mode. These basic functions will also be common to all the other modes. For coded data, the DSG will accept the encoded symbols from some external encoder, e.g., the Test Support Assembly (TSA) in the SPC environment as shown in Fig. 4.

Figure 4 also shows the other DSG components: a SNR generator box (SGB) being built by the Radio Frequency and Microwave Subsystems Section (Section 333), a personal computer (PC) or workstation, and a synthesizer—

the latter two being off-the-shelf. The SGB generates the different signal spectra, including IF, dual subcarrier, or quadrature phase-shift key (QPSK). The monitor and control, as well as other signal-processing programs, will reside in the PC or workstation. Commands will be transmitted to the SNR generator box CPU (SGB-CPU), which in turn will configure the special high-speed hardware needed to generate the output calibrated signals. A programmable frequency synthesizer will generate the variable system clock needed by the high-speed hardware. When the DSG is installed in the DSN-SPC environment, an external interface to the TSA may be used, providing the added capability of coded data, simulated Doppler, and other spacecraft-unique parameters which are already available in the TSA.

Figure 5 shows in block diagram form the main functions assigned to each board residing in a Multibus-I chassis in the SGB. Three channels with identical hardware (pattern generator, filter, and attenuator) will be used to generate two baseband-filtered data channels (channels 1 and 2) of subcarrier binary phase-shift keyed data (BPSK) and one channel of lowpass-filtered noise (noise channel) to be subsequently added to produce an analog output

$$S_T(t) = d_1(t)\sin w_{sc1}t + d_2(t)\sin w_{sc2}t + N_{BB}(t) \quad (4)$$

where

$w_{sc1,2}$  = the first or second subcarrier frequency, rad/sec

$d_{1,2}$  = the first or second baseband-filtered data process

$N_{BB}$  = the baseband-filtered noise process

$\sin x = \text{sign}[\sin(x)]$  or  $\sin(x)$

Single channel generation will be attained by elimination of one of the data channels from Eq. (4).

In case of residual carrier generation, channel 1 will generate the carrier, channel 2 will generate the modulation, and the noise channel will generate the bandpass-filtered noise.

Therefore, the filtered output of the DSG in this configuration will be

$$S_T(t) = \cos \Theta(t) \sin w_c t + m(t) \sin \Theta(t) \cos w_c t + N_{BP}(t) \quad (5)$$

where

<sup>1</sup> *Deep Space Network Flight Project Interface Design Handbook*, JPL 810-5, Rev. D (internal document), Jet Propulsion Laboratory, Pasadena, California.

$w_c$  = the carrier frequency, rad/sec

$\Theta$  = the modulation index

$N_{BPF}$  = the bandpass-filtered noise process

$m(t)$  = the modulation signal

If QPSK signal generation is desired, channel 1 will generate the bandpass-filtered in-phase component, channel 2 will generate the bandpass-filtered quadrature component, and the noise channel will generate the required bandpass-filtered noise. The filtered analog output of the DSG will therefore be

$$S_T(t) = d_I(t) \sin w_c t + d_Q(t) \cos w_c t + N_{BPF} t \quad (6)$$

A concise description of the DSG's different functions follows. Refer to Fig. 5 for architectural details.

### A. Pattern Generation

The pattern generator, when configured to generate noise, will accept data bytes (from the SGB-CPU through the Multibus-I interface) to be mapped into a random-access memory (RAM). After the RAM has been loaded, it will be sequentially read by a random-address generator whose output will represent the RAM's address to be accessed. This random-address generator will be a pseudonoise (PN) code generator with inherent uniform output distribution and a very long period (longer than 24 hours). The distribution function of the data bytes mapped into the RAM in conjunction with the uniform distribution function of the addresses with which this RAM is read will determine the distribution function of the noise generator output. Usually the distribution function mapped into the RAM will be a Gaussian or normal distribution quantized to  $B$  bits from the corresponding analog probability function

$$f(x; u, s) = 1/(s\sqrt{2\pi})e^{-1/2}[(x - u)/s]^2 \quad (7)$$

with  $u = 0$  for unbiased noise and  $s$  = standard deviation.

The discrete probability function will therefore be

$$P_d(x_d; u, s) = 1/2\{\text{erf}[(x_d + 0.5 - u)/(\sqrt{2}s)] - \text{erf}[(x_d - 0.5 - u)/(\sqrt{2}s)]\} \quad (8)$$

for  $x_d \in \{-(L - 2), \dots, -1, 0, 1, 2, \dots, (L - 2)\}$ ,

$$P_d(x_d; u, s) = 1/2\{1 + \text{erf}[(x_d + 0.5 - u)/\sqrt{2}/s]\} \quad (9)$$

for  $x_d = -L - 1$ , and

$$P_d(x_d; u, s) = 1/2\{1 + \text{erf}[(x_d - 0.5 - u)/\sqrt{2}/s]\} \quad (10)$$

for  $x_d = L - 1$ , where the subscript  $d$  denotes discrete. In addition, for  $B$  = number of bits including sign

$$L = 2^{(B-1)} \quad (11)$$

$$s < L/3 \quad (12)$$

From [2],

$$1 + \text{erf}[(x - u)/(s\sqrt{2})] =$$

$$1/(s\sqrt{2\pi}) \int_{-\infty}^x e^{-1/2}[(t - u)/s]^2 dt \quad (13)$$

and

$$\text{erf}(-x) = -\text{erf}(x) \quad (14)$$

In the present breadboard design

$$B = 8, \quad L = 128, \quad s < 43, \quad \text{and} \quad u = 0 \quad (15)$$

and the output noise sequence,  $N_0(n)$ , will be normally distributed with statistical parameters defined by the mapping itself and mainly dependent on the uniformity of the random-address generator and not on its autocorrelation function.

The pattern generator may also be programmed to generate a subcarrier frequency and a data pattern. In both cases, special waveforms or encoded (convolutional, Reed-Solomon, etc.) data sequences may also be mapped into either RAM. The possibility of generating very long random data sequences is also available by reading the data RAM with the random address generator, as in the case of generating random noise. The normal configuration will be the sequential reading of the data mapped into the RAM. The time sequences derived by the sequential reading of both data RAM,  $d(n)$ , and subcarrier RAM,  $Sc(n)$ , are multiplied at the system clock rate ( $T_{s,y_s}$ ) to obtain a BPSK modulation sequence

$$D(n) = d(n)Sc(n) \quad (16)$$

It was considered that a data pattern of 65,536 bits would be the maximum length ever to be required; therefore, the RAMs implemented will be 64K RAMs.

## B. Digital Filtering and Attenuation

The main purpose of this digital filter and attenuator circuitry will be to provide a user-defined lowpass filter when generating a baseband spectrum, or a bandpass filter when generating an IF spectrum. Noise and data will be independently filtered by two digital filters. These filters may have, if desired, the same frequency response, in which case the output difference equation [3] will be for the data path:

$$D_f(n) = h(0)D(n) + h(1)D(n-1) + \dots + h(N-1)D(n-N+1) \quad (17)$$

and for the noise path

$$N_{0f}(n) = h(0)N_{0f}(n) + h(1)N_{0f}(n-1) + \dots + h(N-1)N_{0f}(n-N+1) \quad (18)$$

with  $h(n)$  = finite impulse response (FIR).

A 63-tap FIR filter ( $N = 63$ ) will be implemented. The user may define the filtering process independently for each noise or data path by simply modifying the coefficients  $h(n)$ . Filter coefficients' definition will be attained by DSP software residing on the workstation or PC (refer to Fig. 4). After independent filtering, the noise and data outputs are properly scaled (attenuated) and added to generate the desired symbol SNR on the output sequence

$$S_T(n) = A_D D_f(n) + A_N N_{0f}(n) \quad (19)$$

with  $A_D \leq 1$  for the signal path attenuator factor and  $A_N \leq 1$  for the noise path attenuator factor. Note that, due to this independent filtering and individual attenuation, the filtered data spectrum will be known. This precise knowledge provides the basis for optimum subcarrier and data demodulation processes [4] to be performed later on the statistics function. The DSG SNR output, or equivalently the SNR input, to the demodulator under test will, thus, be continuously monitored by this optimum process.

## C. Digital-to-Analog Conversion

The DSG will provide a digital output for baseband synchronous testing and, through the digital-to-analog conversion function, an analog output for a more general asynchronous type of testing. Therefore, the input sequence Eq. (19) containing the subcarrier, data, and noise at the

system clock rate is converted to the analog representation in Eqs. (4), (5), or (6) by a digital-to-analog converter (DAC). This DAC is the element dictating the number of bits to be used in the DSG design. Bandwidth requirements limited to 12 the number of bits to be used at the time of design definition and will be the quantization used in the prototype demonstration. The DAC analog output spectrum will be rich in harmonics that have to be attenuated by an output analog filter. The amplitude, and more important, the phase characteristics of this filter have to be very well controlled in the design in order to conserve the input spectrum characteristics. The effect on SNR of several Butterworth filters with numbers of poles ranging from three to nine were simulated with the conclusion that in the worst case, to obtain less than 0.1-dB output SNR degradations, a three-pole Butterworth filter should be used.

## D. Statistics Monitor

The DSG will be used to calibrate SNR measurements and losses on other signal processes; therefore, several statistical measurements have been implemented through the statistics' monitor function. The same statistical measurements will be implemented in the digital output, as well as in the analog output. Note that the digital output time sequence Eq. (19) is directly brought to the statistics board where an optimum subcarrier and data demodulation is performed by digital multiplication of this time series by an exact replica of the data spectrum embedded in that output

$$S(n) = S_T(n)A_D D_f(n) \quad (20)$$

where the delay of the digital hardware has been arbitrarily set to zero due to its precise knowledge.

The result of that multiplication is accumulated in the symbol integrator and dump for exactly a symbol period ( $T_S$ ), related to the system clock period ( $T_{sys}$ ) by the relationship

$$T_S = I_S T_{sys} \quad (21)$$

where  $I_S$  is an integer.

Therefore the  $i$ th symbol integrated value will be

$$S_i = \sum_{n=(i-1)I_S+1}^{iI_S} S(n)$$

$$= \sum_{n=(i-1)I_S+1}^{iI_S} [A_D^2 D_f^2(n) + A_D A_N N_{0f}(n) D_f(n)] \quad (22)$$

Note that all statistics are performed in this integrated symbol output,  $S_i$ . Thus, the mean value of the detected symbols will be calculated on the SGB-CPU from results of the symbol value accumulator obtained for a fixed number of symbol periods ( $K$ ) completing approximately one second of elapsed time

$$\bar{S} = 1/K \sum_{i=1}^K S_i \quad (23)$$

Simultaneously, the squared value of the detected symbol is also accumulated in the symbol squared accumulator for the same number ( $K$ ) of symbols

$$\bar{S}^2 = 1/K \sum_{i=1}^K S_i^2 \quad (24)$$

From these two values the symbol SNR evaluation immediately follows (in decibels)

$$SNR_M = 10 \log_{10} \left\{ \bar{S} / \left[ 2(\bar{S}^2 - \bar{S}^2) \right] \right\} \quad (25)$$

The last measurement made in this output is the symbol error count in the symbol error accumulator. This measurement is arrived at by just counting the output negative events in the same period of time ( $KT_{sy}I_S$ ).

$$SER = 1/K \sum_{i=1}^{K+i} \text{neg}[S_i] \quad (26)$$

with

$$\begin{aligned} \text{neg}[S_i] &= 1 \text{ if } S_i < 0 \\ \text{neg}[S_i] &= 0 \text{ if } S_i > 0 \end{aligned} \quad (27)$$

In order to characterize the hardware performance, a histogram accumulator will be implemented, thereby providing a straightforward method to confirm the actual probability density function of the filtered or unfiltered noise from Eq. (18). Statistics on data and subcarrier waveforms can be performed and will be used as a hardware-software performance verification self test.

To calibrate the analog output SNR, the first function to be performed is the digital conversion of the analog output. In the case of baseband or IF testing, the analog-to-digital conversion is performed directly on the DSG analog output through a 12-bit analog-to-digital convertor. In the case of higher frequency spectra, a downconversion will precede the analog-to-digital conversion, as shown previously in Fig. 2. In both cases, the delay introduced by the analog circuitry,  $T$ , has to be accounted for. This delay ( $T$ ) will not necessarily be an integer number of system clock cycles ( $k$ ), but rather will have also a fractional part ( $\tau$ )

$$T = kT_{sys} + \tau$$

Therefore, after removing the integral part of system clock cycles,  $kT_{sys}$ , Eq. (22) will now be

$$\begin{aligned} S_i &= \sum_{n=(i-1)I_S+1}^{iI_S} S(n) \\ &= \sum_{n=(i-1)I_S+1}^{iI_S} [A_D^2 D_f(n) D_f(n + \tau) \\ &\quad + A_D A_N N_{0f}(n) D_f(n + \tau)] \end{aligned} \quad (28)$$

To account for this unknown fractional delay,  $\tau$ , a delay line of 2-nsec steps will be implemented. SNR degradation due to these quantization steps will be (in decibels)

$$\Delta SNR = 20 \log_{10}(1 - 2\tau/T_S) \quad (29)$$

with  $\tau$  = delay quantization and  $T_S$  = symbol period. In any case, this degradation will be calibrated through the autocorrelation function Eq. (28) obtained with 2-nsec quantized steps.

### III. Error Analysis

Whenever practical, the error contributions to the SNR generation were evaluated by analysis or computer simulation (A). If impractical, an engineering judgment approach (EJ) was taken. The expected errors on the setability and the actual measurement of the DSG output SNR are summarized in Tables 1 and 2. The total error contribution will be verified when the design is completed. Individual errors will also be measured whenever possible.

The errors to be encountered when a given SNR is desired in the analog output, assuming a memoryless DSG,

have been summarized in Table 1. Note that subsequent SNR settings may be known with the accuracy deduced in Table 2 when the operator takes advantage of the accuracy provided by the SNR monitor (the operator has now provided memory to the process).

The digital output monitoring error will be mainly produced by the deviation from a perfect uniform distribution of the PN generator addressing the Gaussian noise RAM and the quantization noise of the digital process. Both have been quantified in Table 1 for one-second integration times with an evaluated error of 0.01 dB for the nonuniformity of the random-address generator. This error will diminish as the integration time is increased, with a limit imposed by the quantization noise (0.002 dB). Therefore, the accuracy of the digital SNR measurement will only be dependent on the integration time used to obtain that measurement, or equivalently on the number of samples used in the computation with a lower limit imposed by the quantization noise of the digital process—in the present case 0.002 dB. Confidence intervals can be found if the cumulative distribution function of the SNR measurement is known. It has been shown [5] that the distribution function of the SNR measurement, Eq. (25), is a noncentral  $F$ -distribution. The cumulative distribution of this function may be approximated by the standard normal distribution function as follows:

$$F(SNR_M) = P(x) = 1/\sqrt{2\pi} \int_{-\infty}^x e^{-1/2t^2} dt \quad (30)$$

with

$$x = (n_1 - n_2)/(d_1 + d_2)^{1/2} \quad (31)$$

$$n_1 = [(N - 1)SNR_M/(1 + N SNR_T)]^{1/3} \times \{1 - 2/[9(N - 1)]\} \quad (32)$$

$$n_2 = 1 - \left\{ (2 + 4N SNR_T)/[9(1 + N SNR_T)^2] \right\} \quad (33)$$

$$d_1 = (2 + 4N SNR_T)/[9(1 + N SNR_T)^2] \quad (34)$$

$$d_2 = 2/[9(N - 1)][(N - 1)SNR_M/(1 + N SNR_T)]^{2/3} \quad (35)$$

and

$N$  = the number of samples in the measurement

$SNR_T$  = the true SNR

$SNR_M$  = the measured SNR

Figure 6 is a plot of Eq. (30). It gives the cumulative probability distribution of measuring the SNR ( $SNR_M$ ) within 0.1 dB of the true SNR ( $SNR_T$ ) as a function of the number of samples used in the measurement.  $SNR_T$  has been used as the parameter.

Table 2 summarizes the error budget for the SNR measurement on the analog output.

#### IV. Capabilities

Table 3 summarizes the capabilities of the present breadboard design. When the DSG is configured to generate a baseband spectrum, it will be able to provide two data channels with or without subcarrier, or two biphasic data channels in a total baseband bandwidth of 10 MHz (dc to 10 MHz). If the DSG is configured to generate a bandpass spectrum it will be able to provide a residual carrier up to 5 MHz with one data channel with or without subcarrier, or one QPSK or one offset QPSK (OQPSK) data channel, or two data channels (carrier suppressed) in a total bandpass bandwidth of  $\pm 5$  MHz.

The frequency generation on the DSG is related to the system clock rate ( $f_{sys}$ ) in use. The subcarrier frequency ( $f_{sc}$ ) or intermediate frequency (IF) will be generated with a variable frequency resolution ranging from  $f_{sys}/2^{17}$  for the highest frequencies (5 MHz) to  $f_{sys}/2^{32}$  for the lowest (100 Hz). The data period ( $T_S$ ) will be related to the system clock period ( $T_{sys}$ ) by

$$T_{sys} = T_S/I_S \quad (36)$$

with

$$2^1 \leq I_S \leq 2^{24} \quad (37)$$

and

$$T_{sys} = \text{system clock period} = 1/f_{sys}$$

$$T_S = \text{symbol period}$$

$$2 \text{ MHz} < (1/T_{sys} = f_{sys}) < 20 \text{ MHz}$$

An external DSG mode is provided. In this mode, the data waveform (or the subcarrier biphase modulated by the data) is externally generated and sampled at the system clock rate  $f_{sys}$ . In general, the frequencies involved will not be coherently related. Note that the mean frequency derived from this, in general, asynchronous sampling process will still correspond to the data rate (or subcarrier and data) required. Because of knowledge of the asynchronous sampled data spectrum, the DSG SNR-monitoring process will provide an optimum demodulation independent of the sampling process. This will not generally be the case of the demodulation process under test when the symbol clock is synchronous with the incoming symbol or with the mean symbol rate. In this case, the synchronous detection will expect symbols with exactly the same period and not the jittery ones being provided. The worst-case degradation of such a process will be quantified as follows.

Assume that a periodic signal ( $f_{sc}$ ) is sampled at a fixed sampling rate ( $f_{sys}$ ). The resulting frequency spectrum is composed of spectral lines ( $f_H$ ) related to the sampling frequency by

$$f_H = |nf_{sc} - mf_{sys}| \quad (38)$$

with  $\underline{n}$  and  $\underline{m}$  integers.

Assuming, for simplicity, that the sampled periodic signal ( $f_{sc}$ ) is a square wave, the useful spectral lines in a subsequent symbol synchronous demodulation process will only be those corresponding to the odd harmonics of the corresponding Fourier series with corresponding total power

$$P_{sq} = \sum_i^{\infty} [2A/(i\pi)]^2 \quad (39)$$

for ( $i = 1, 3, 5, \dots$ ). The remaining alias harmonics given by Eq. (38) may, in the most general sense, be treated as unwanted noise in the subsequent symbol synchronous demodulation process. Given a Nyquist bandwidth, if the useful square-wave signal power is

$$P_{sqN} = \sum_i^N [2A/(i\pi)]^2 \quad (40)$$

for ( $i = 1, 3, 5, \dots, N$ ) with

$$Nf_{sc} \leq 1/2f_{sys} \quad (41)$$

and the total available power is  $P_T$ , then the available SNR on the square wave (or subcarrier SNR) is defined as

$$SNR_{sqN} = 10 \log_{10}[P_{sqN}/(P_T - P_{sqN})] \quad (42)$$

The above process was computer simulated for 20 equally spaced subcarrier frequencies (0.0977 MHz to 4.7363 MHz) and a system clock of  $f_{sys} = 20$  MHz. Results are shown in Fig. 7. Note that discontinuities will exist due to the changing number of odd harmonics in the ideal rectangular Nyquist filter.

Other discontinuities will exist at subcarrier frequencies ( $f_{sc}$ ) corresponding to an exact even submultiple (2, 4, 6, ...) of the system clock rate ( $f_{sys}$ ). In this case, the sampling process becomes synchronous and the available SNR becomes infinity ( $\infty$ ) at those particular frequencies. This synchronous mode will have to be used for calibrations of data rates higher than 2 MS/sec (1-MHz square wave). All the required operations in the external mode will be transparent to the DSG operator and will be automatically software-controlled.

## V. Conclusions

A digital synthesis method to generate bandlimited test signals with precise signal-to-noise ratios has been described and the expected errors have been quantified. Design of appropriate hardware and software to demonstrate the performance of this method has been initiated. The DSG is expected to replace the presently used Y-factor method with improvements on testing time, accuracy, stability, and repeatability.

## Acknowledgments

The author wishes to thank B. Parham, C. Chang, L. Howard, and M. Marina for their contributions to the detailed design and enthusiastic dedication to the breadboard construction of the proposed SNR generator. The author also thanks J. Statman and T. Peng for suggestions in the process of reviewing this article.

## References

- [1] W. S. Baumgartner, N. C. Ham, W. F. McAndrew, D. W. Brown, M. L. Yeater, C. A. Holritz, J. T. Hatch, and A. D'Amore, "Multiple-Mission Telemetry System Project," *JPL Space Programs Summary 37-60*, vol. II, pp. 152-169, August 1970.
- [2] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Applied Mathematics Series, Washington, D.C.: National Bureau of Standards, December 1972.
- [3] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, New Jersey: Prentice-Hall, Inc., 1975.
- [4] R. Sadar and W. J. Hurd, "Detection of Signals by the Integrate-and-Dump Filter With Offset Sampling," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 158-173, November 15, 1987.
- [5] W. H. Tranter and M. D. Turner, "Signal to Noise Ratio Estimation in Digital Computer Simulation of Lowpass and Bandpass Systems With Applications to Analog and Digital Communications," *NASA Contract No. NAS 9-14848*, Missouri: University of Missouri at Rolla, July 1977.

**Table 1. Setability error budget (worst case)**

Error	dB	Source
RAM (64K, 8-bit) resolution	0.04	A
PN generator uniformity	0.01	EJ
Quantization noise	0.002	A
DAC nonlinearities	0.01	A
DAC frequency response	0.01	A
System clock phase jitter	0.02	EJ
Analog filter	0.1	A
Amplifiers	0.01	EJ
Total error on SNR setting (root sum square, rss)	0.11	

Note: Evaluated by: (A) analysis or computer simulation  
(EJ) engineering judgment

**Table 2. Analog output monitoring error budget**

Error	dB	Source
Quantization noise	0.002	A
System clock phase jitter	0.02	EJ
ADC quantization noise	0.002	A
ADC nonlinearities	0.01	A
Waveform distortion	0.01	A
Total error on SNR monitor (rss)	0.02	

Note: Evaluated by: (A) analysis or computer simulation  
(EJ) engineering judgment

**Table 3. Summary of DSG capabilities**

Function	Capability
System clock	2 to 20 MHz
Analog power output	7 dBm (50 ohms)
Noise density distribution	Gaussian
Noise bandwidth	0.1 to 10 MHz
Noise autocorrelation period	> 24 hr
Baseband external input	TSA compatible
Symbol SNR	
Setability	0.1 dB
Accuracy and stability	±0.05 dB
Range	-6 to 20 dB
Data types	Nonreturn to zero; biphase; QPSK; OQPSK
Data rates	4 S/sec to 6.6 MS/sec
Modulation types	Phase-shift keyed
Data pattern	2048 and 16,384 PN code
Data transition density	10; 30; 50; 70; and 100%
Frame synchronization word	Up to 64 bits
Frame length	Up to 65,536 bits
Subcarrier or IF frequencies	100 Hz to 5 MHz

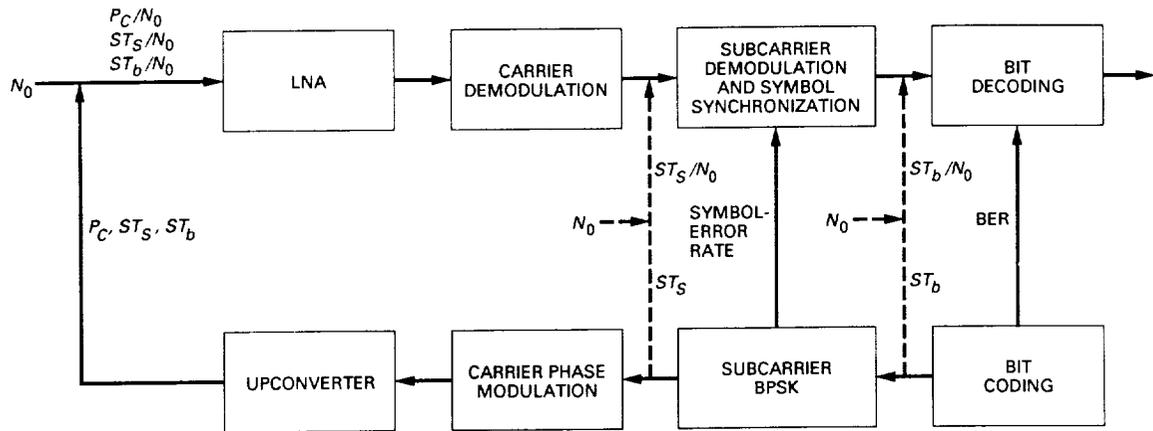


Fig. 1. The SPC portion of a typical DSN communications link.

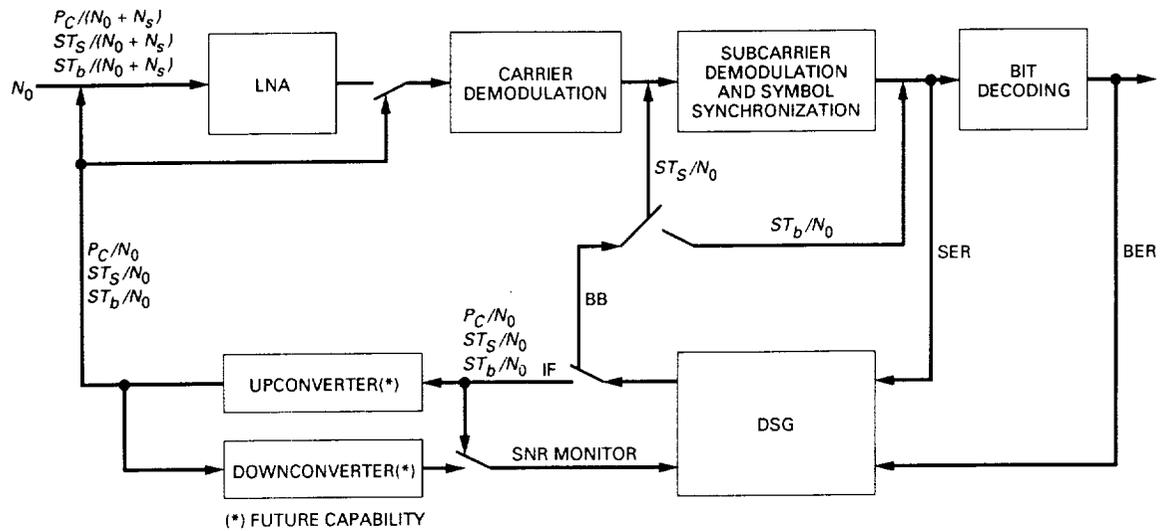


Fig. 2. Calibration of SPC system degradation and losses.

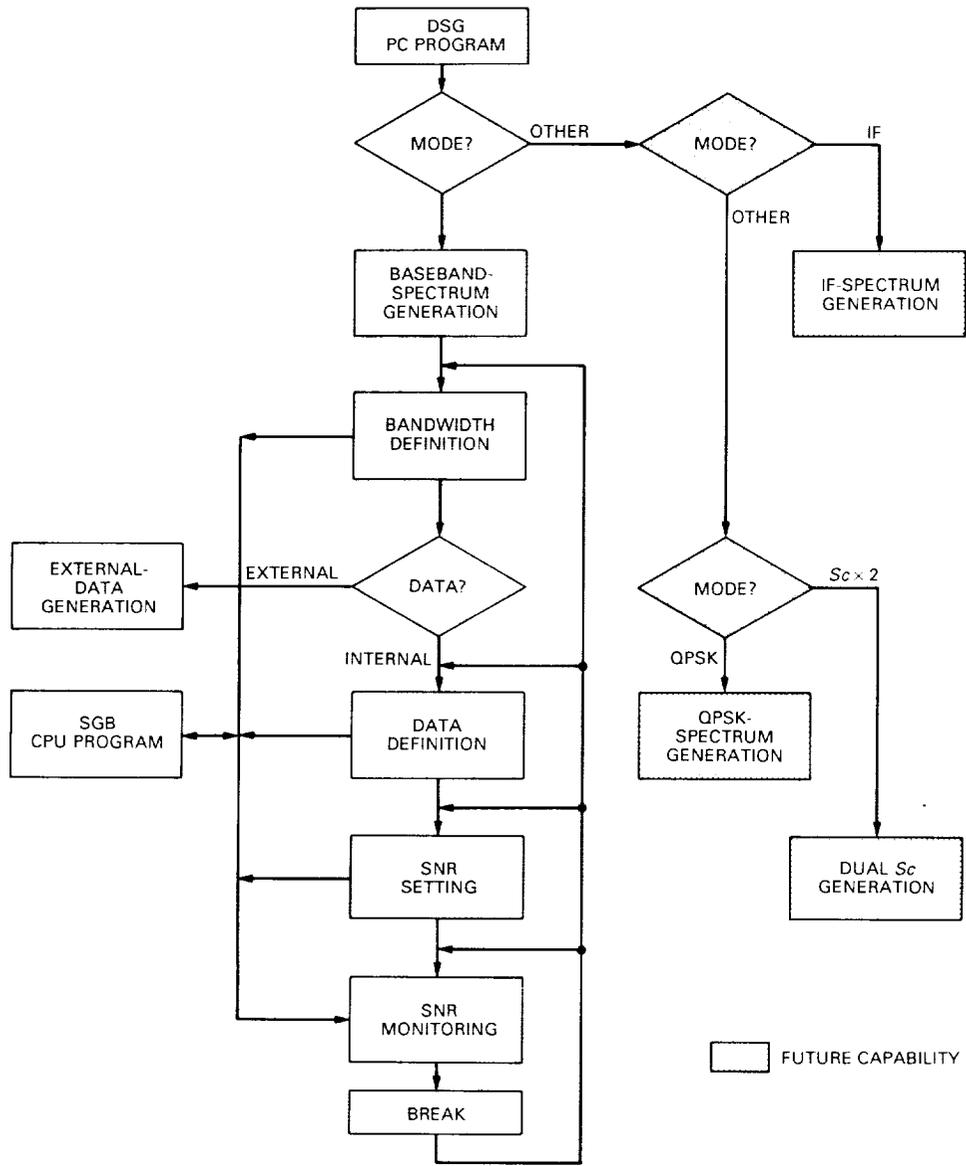


Fig. 3. Programmable capabilities of the DSG.

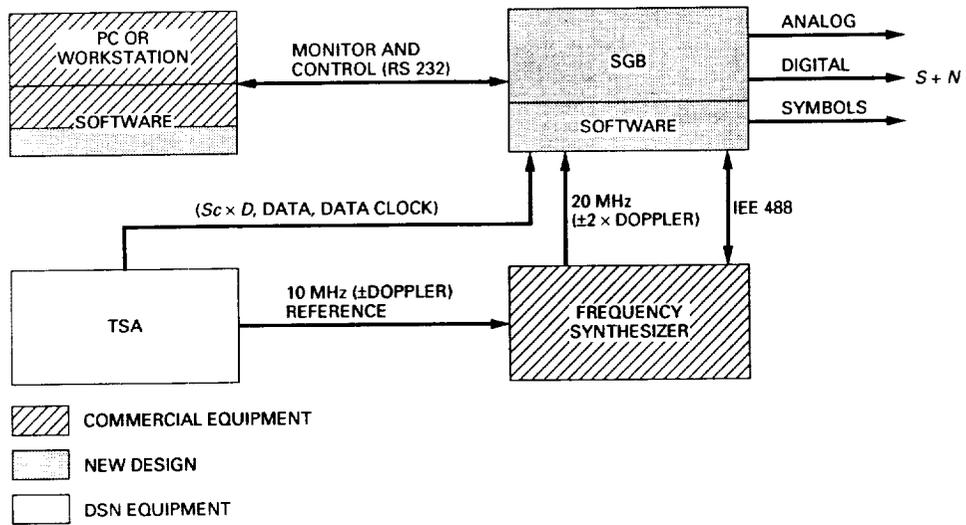


Fig. 4. The DSG external interfaces in the SPC environment.

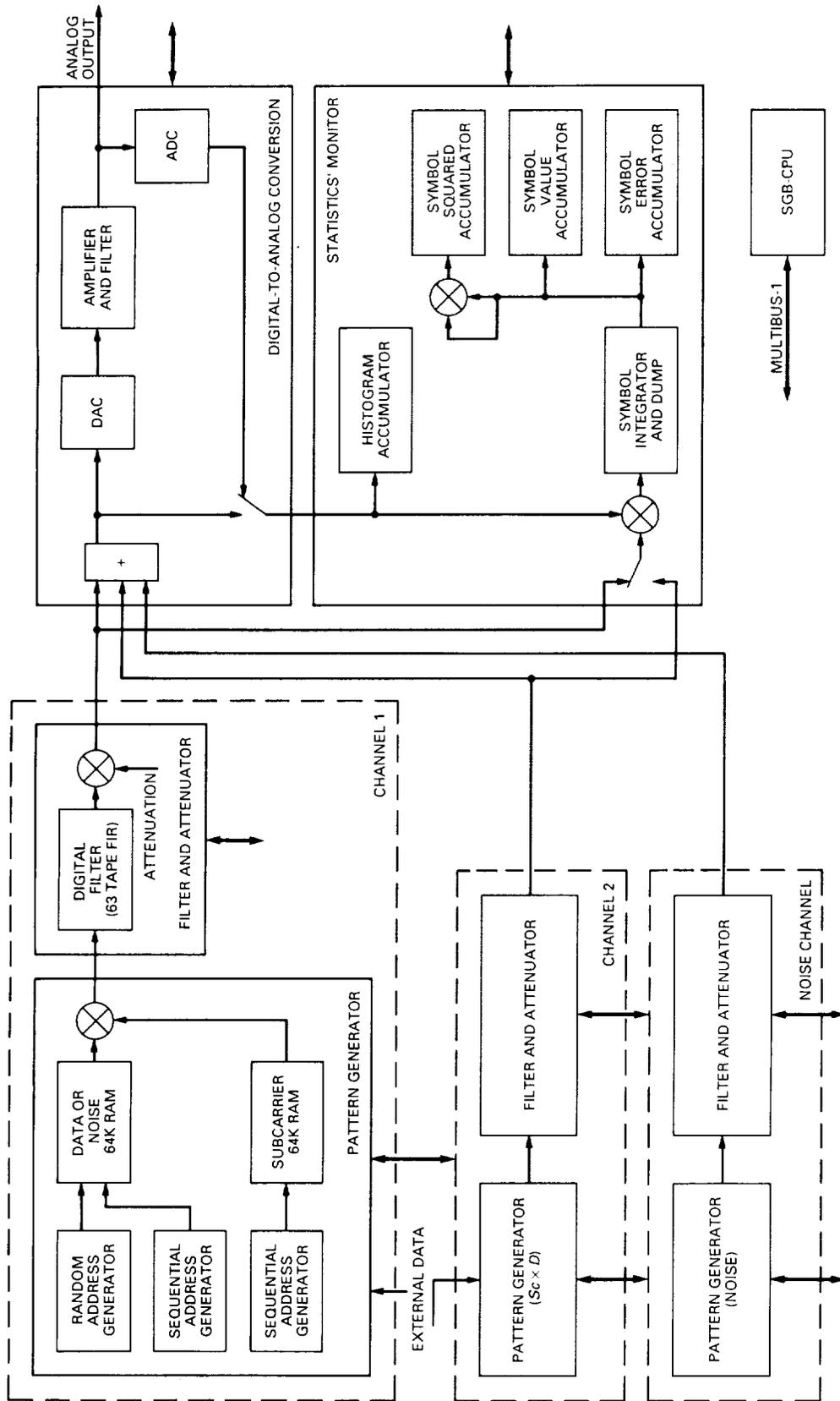


Fig. 5. SNR generator box.

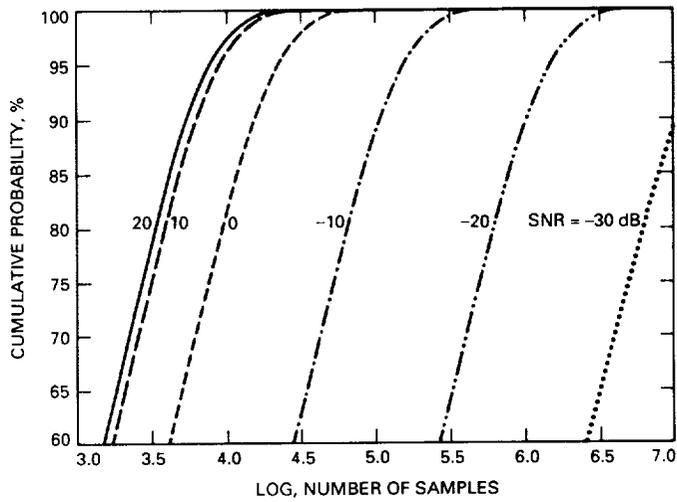


Fig. 6. Cumulative probability (of measured SNR being within 0.1 dB of true SNR) versus total number of samples.

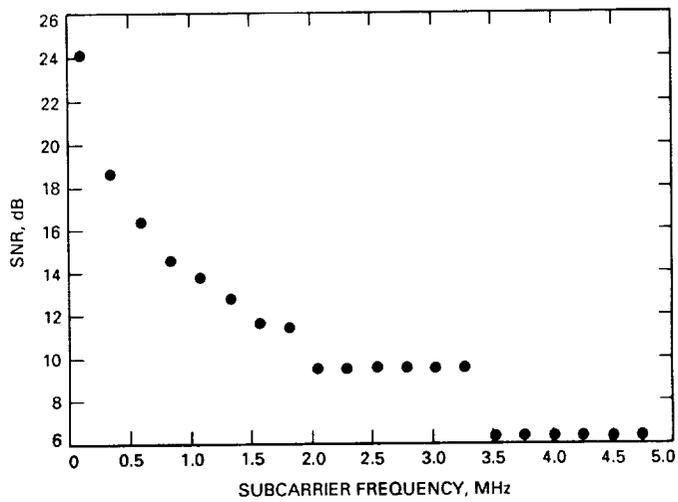


Fig. 7. Upper limit of subcarrier SNR (due to asynchronous sampling) versus subcarrier frequency.

S/S-32  
332013

P-17

**N91-18323**

JJ5-8490

# A Comparison of Manchester Symbol Tracking Loops for Block V Applications

J. K. Holmes

Radio Frequency and Microwave Subsystems Section

*The linearized tracking errors of three Manchester (biphase-coded) symbol tracking loops are compared to determine which is appropriate for Block V receiver applications. The first is a nonreturn-to-zero (NRZ) symbol synchronizer loop operating at twice the symbol rate ( $NRZ \times 2$ ) so that it operates on half-symbols. The second near optimally processes the mid-symbol transitions and ignores the between-symbol transitions. In the third configuration, the first two approaches are combined as a hybrid to produce the best performance. Although this hybrid loop is the best at low symbol signal-to-noise ratios (SNRs), it has about the same performance as the  $NRZ \times 2$  loop at higher SNRs ( $> 0\text{-dB } E_S/N_0$ ). Based on this analysis, it is tentatively recommended that the hybrid loop be implemented for Manchester data in the Block V receiver. However, the high data rate case and the hardware implications of each implementation must be understood and analyzed before the hybrid loop is recommended unconditionally.*

## I. Introduction

Three symbol-synchronization (sync) loops have been studied with the object of determining which structure provides the best tracking performance in terms of the minimum tracking error variance of the linearized loop:

- (1) The nonreturn-to-zero (NRZ) digital data transition tracking loop (DTTL), which operates at twice the Manchester symbol rate (or at the equivalent NRZ symbol rate).
- (2) A symbol-sync loop based on a near-optimal processing of the mid-symbol transition. The between-symbol transitions are ignored by this loop.
- (3) A hybrid of loops (1) and (2). The mid-symbol transition processing is based on the second candidate

loop and the between-symbol transition processing is based on the DTTL, in which a transition is estimated from the half-symbol on either side of the between-symbol transition.

Other possibilities exist, but these three seemed most relevant and more readily analyzable.

To make the analysis somewhat simpler to accomplish, the assumption was made that the symbol tracking loops are continuous in time and amplitude. Thus, the results given here would apply to the Block V digital receiver only at the low and medium symbol rate cases, and not to the high symbol rate case where as few as three samples can occur per half-symbol.

## II. Analysis of DTTL Tracking Error Variance Operating at $2R_S$

In this section the continuous-time linearized closed-loop tracking performance (expressed in fractions of symbol time) is estimated. The symbol tracking loop under consideration is an NRZ DTTL, which works on the NRZ symbol transitions to detect the timing error. Figure 1 shows a block diagram of the NRZ DTTL operating at twice the symbol rate ( $2R_S$ ) so that it is suitable for tracking Manchester (biphase) signal formatting. Note that because false lock can occur on any Manchester-coded symbol tracking loop, a false-lock detector must be used with Manchester data.

In Fig. 1,  $\hat{\tau}$  denotes the symbol loop estimate of the symbol stream transmission delay and  $W$  is the window size in seconds used for the error-detection window.

Basically the loop performs one integration over one complete half-symbol ( $T_H$  sec) and another across the time where the transition could occur. When the transition is mid-symbol, a transition always occurs; when the transition is at the end of the symbol, a transition may or may not occur.

The input is modeled as an infinite sequence of Manchester symbols with transitions determined by the estimate of the half-symbol sequence  $a_k$ . In addition, the thermal noise corrupts the symbol stream.

Thus, the received signal is modeled as

$$y(t) = \sqrt{P} \sum_{k=-\infty}^{\infty} b_k q(t - kT - \tau) + n(t) \quad (1)$$

where

$q(t)$  is one Manchester symbol

$n(t)$  is modeled as white Gaussian noise (WGN) with spectral density  $N_0/2$

$b_k$  is a random binary valued ( $\pm 1$ ) symbol sequence

$T_H$  is the symbol half-period ( $T = 2T_H$ )

$T$  is the symbol period

$P$  is the data power in the received signal

$\tau$  is the time delay of the signal

$\hat{\tau}$  is the time-delay estimate of the symbol sync loop

The relationship between the half-symbol  $a_k$  and full-symbol data  $b_k$  sequence is given by

$$\left. \begin{aligned} a_k &= b_{k/2}, & k \text{ even} \\ a_k &= -b_{k/2-1/2}, & k \text{ odd} \end{aligned} \right\} \quad (2)$$

and is illustrated in Fig. 2. The general approach used in [1] is followed and it is assumed for convenience that  $\tau = 0$ .

The inphase channel produces an output pulse sequence estimate, which, for  $\tau - \hat{\tau}$  close to zero, is essentially given by

$$\hat{a}_k = \text{SGN} \left( a_k \sqrt{P} T_H + \int_{kT_H}^{(k+1)T_H} n(t) dt \right), \quad (\tau - \hat{\tau} \cong 0) \quad (3)$$

where  $k$  is the index on half-symbols. The output of the transition indicator is given by

$$I_k = \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \quad (4)$$

At the end of the  $k$ th pulse time, the midphase channel produces the following output<sup>1</sup> when  $\tau - \hat{\tau} \cong 0$  (and  $\tau = 0$ ):

$$U_k = 2a_{k-1} \sqrt{P} (\tau - \hat{\tau}) + \int_{kT_H - W/2}^{kT_H + W/2} n(t) dt, \quad |\tau - \hat{\tau}| \leq W/2 \quad (5)$$

Consequently, the timing-error estimate  $\hat{\epsilon}$ , which is the loop estimate of the timing error  $\epsilon = \tau - \hat{\tau}$ , is given by

$$\hat{\epsilon}(t) = \sum_{k=-\infty}^{\infty} I_k U'_k p(t - (k+1)T_H) \quad (6)$$

which changes every  $T_H$  seconds. Now  $p(t)$  is a half-Manchester pulse of unity amplitude (see Fig. 8) and  $U'_k$  is the  $(T_H - W/2)$ -sec delayed version of  $U_k$ . It is used to align the midphase and inphase channels in time. This error signal is constant over  $T_H$  sec in Eq. (6). Using Eqs. (4) and (5) in Eq. (6), the expression for the loop error signal is obtained:

<sup>1</sup> The timing error is neglected in the noise term but included in the error-signal term.

$$\begin{aligned} \hat{\varepsilon}(t) = & \sum_{k=-\infty}^{\infty} \left\{ a_k \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right) 2\sqrt{P}(\tau - \hat{\tau}) \right. \\ & \left. + \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right) \int_{kT_H - W/2}^{kT_H + W/2} n(t) dt \right\} \\ & \times p(t - (k+1)T_H) \end{aligned} \quad (7)$$

It will be shown that the mean value of  $\hat{\varepsilon}$  is given by  $\alpha\varepsilon$  (linear) for small values of  $\varepsilon$ ; the resulting additive noise process is denoted by  $N(t)$ . Both can be determined from Eq. (7) since the noise is a random amplitude pulse sequence process. It is assumed that  $\varepsilon$  is small in the following discussion. Then the timing estimate  $\hat{\tau}$  is given by

$$\hat{\tau} \cong \frac{KF(s)}{s} [\alpha\varepsilon + N(t)] \quad (8)$$

where  $F(s)$  is the loop filter expressed in Heaviside operator symbolism  $(1/s)X(s)$  denotes  $\int_0^t x(t')dt'$  and represents the effect of the voltage-controlled oscillator (VCO).

Since by definition of the error  $\varepsilon$

$$\hat{\tau} = \tau - \varepsilon \quad (9)$$

Using Eq. (9) in Eq. (8) yields

$$\varepsilon(t) = \left( \frac{K'F(s)/s}{1 + K'F(s)/s} \right) \left( \frac{N(t)}{\alpha} \right) \quad (10)$$

where  $F(s)$  is the loop filter function viewed as a Heaviside operator and the  $1/s$  comes from the VCO. The terms that depend upon  $s$  comprise the closed-loop transfer function; it is denoted by  $H(s)$ , so that Eq. (10) becomes

$$\varepsilon(t) = H(s) \left( \frac{N(t)}{\alpha} \right) \quad (11)$$

where again  $H(s)$  is viewed as a Heaviside operator operating on the noise term following it.

Next it is necessary to characterize the noise process  $N(t)$  and the constant  $\alpha$ . First consider the computation of  $E[\hat{\varepsilon}|\varepsilon]$ . For small timing errors, it will be assumed that the value of  $\hat{a}_k$  is statistically independent of the integrated

noise process (integrated from  $kT_H - W/2$  to  $kT_H + W/2$ ). Of course, this is not true but it has been demonstrated by simulation to be a reasonable approximation [2]. With this assumption, Eq. (7) can be used to obtain

$$\begin{aligned} E[\hat{\varepsilon}|\varepsilon] \cong & 2\sqrt{P}(\varepsilon) \left[ \frac{1}{2} (1) \left\{ \overbrace{(1 - PE_H)^2 - PE_H^2}^{\text{mid-symbol transition}} \right\} \right. \\ & \left. + \frac{1}{2} \left( \frac{1}{2} \right) \left\{ \overbrace{(1 - PE_H)^2 - PE_H^2}^{\text{adjacent symbol transition}} \right\} \right] \end{aligned} \quad (12)$$

where the two leftmost  $1/2$  factors in the rectangular brackets are due to the probability of the transition being a mid-symbol transition or an adjacent symbol transition. The factor of unity following the first factor of  $1/2$  accounts for the fact that there is always a mid-symbol transition. The factor of  $1/2$  following the second leftmost factor of  $1/2$  is based on the assumption that there is a probability of  $1/2$  that there is a transition at the end of the symbol. Finally,  $PE_H$  is the probability of a half-symbol error and is given by

$$PE_H = \int_{\sqrt{2R_H}}^{\infty} \frac{1}{\sqrt{2R_H}} e^{-z^2/2} dz = Q(\sqrt{2R_H}) \quad (13)$$

where

$$R_H = \frac{E_H}{N_0} = \frac{PT_H}{N_0} \quad (14)$$

where

$P$  is the data power

$N_0$  is the one-sided noise spectral density at the symbol sync input

$T_H$  is one-half the symbol duration

Hence, from Eqs. (12) and (13)

$$E[\hat{\varepsilon}|\varepsilon] \cong \frac{3}{2} \sqrt{P} \varepsilon \left( 1 - 2Q\sqrt{2R_H} \right) \quad (15)$$

This can be rewritten as

$$E[\hat{\varepsilon}|\varepsilon] = \frac{3}{2}\sqrt{P}\varepsilon \operatorname{erf}(\sqrt{R_H}) = \frac{3}{2}\sqrt{P}\varepsilon \left( 2 \int_0^{\sqrt{R_H}} \frac{1}{\sqrt{\pi}} e^{-t^2} dt \right) \quad (16)$$

Thus,  $\alpha$  of Eq. (8) is given by ( $\alpha =$  the slope of the  $S$ -curve at  $\varepsilon = 0$ )

$$\alpha = \frac{3}{2}\sqrt{P} \operatorname{erf}(\sqrt{R_H}) \quad (17)$$

Now the noise spectral density of  $N_E(t)$  is obtained from the process generated by

$$N_E(t) = \sum_{k=-\infty}^{\infty} \int_{kT_H - W/2}^{kT_H + W/2} n(t) dt \left[ \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right] \times p(t - (k+1)T_H) \quad (18)$$

where, as before,  $p(t)$  is a unit amplitude pulse of duration  $T_H$  sec long. Again assume that  $\hat{a}_k$  is independent of  $n(t)$ ,

and note that the cyclostationary process  $N_E(t)$  can be made stationary by averaging over time [3]. Thus

$$R(\zeta) = \frac{1}{T} \int_0^T E[N_E(t)N_E(t+\zeta)] dt \quad (19)$$

is the autocorrelation function of a stationary process derived from the corresponding cyclostationary process. An evaluation of Eq. (19) obtains

$$R(\zeta) = \frac{N_0 W}{2} \left( 1 - \frac{|\zeta|}{T_H} \right) E \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right)^2 \text{ for } |\zeta| \leq T_H$$

$$= 0 \text{ elsewhere} \quad (20)$$

where it is assumed for analytic convenience that the noise process  $n(t)$  over  $W$  sec is statistically independent of the symbol estimate. Consider the term inside the expectation. It can be evaluated by

$$E \left[ \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right)^2 \right] = \frac{1}{2} E \left[ \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right)^2 \Big|_{\text{mid-symbol transition}} \right] + \frac{1}{2} E \left[ \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right)^2 \Big|_{\text{adjacent-symbol transition}} \right] \quad (21)$$

An evaluation obtains

$$E \left[ \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right)^2 \right] = \frac{1}{2} \left[ \overbrace{1 \cdot (1 - PE_H)^2 + 1 \cdot PE_H^2}^{\text{mid-symbol transition}} \right] + \frac{1}{4} \left[ \overbrace{1 \cdot 2PE_H(1 - PE_H) + (1 - PE_H)^2 + PE_H^2}^{\text{adjacent symbol transition}} \right] \quad (22)$$

or, simplifying

$$E \left[ \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right)^2 \right] = \frac{3}{4} - PE_H(1 - PE_H) \quad (23)$$

Note that when  $PE_H \rightarrow 0$ , the expectation approaches  $3/4$  as anticipated, and when  $PE_H \rightarrow 1/2$ , the expectation approaches  $1/2$  as anticipated. Since

$$1 - 2PE_H = \operatorname{erf}(\sqrt{R_H}) \quad (24)$$

one obtains

$$E \left[ \left( \frac{\hat{a}_k - \hat{a}_{k-1}}{2} \right)^2 \right] = \frac{3}{4} - \frac{1}{4} \left( 1 - \operatorname{erf}^2(\sqrt{R_H}) \right) \quad (25)$$

Therefore

$$R(\zeta) \cong \frac{N_0 W}{2} \left(1 - \frac{|\zeta|}{T_H}\right) \left\{ \frac{3}{4} - \frac{1}{4} \left(1 - \operatorname{erf}^2(\sqrt{R_H})\right) \right\}$$

for  $|\zeta| \leq T_H$

$$= 0 \quad \text{elsewhere} \quad (26)$$

Integrating  $R(\zeta)$  from  $-\infty$  to  $\infty$  yields the noise spectral density at  $f = 0$ , which is

$$\mathcal{S}_N(0) = \left(\frac{N_0 W}{2}\right) \frac{3}{4} \left[1 - \frac{1}{3} \left(1 - \operatorname{erf}^2(\sqrt{R_H})\right)\right] T_H \quad (27)$$

Thus, the absolute, linearized, tracking-error variance is obtained from Eq. (11) to be

$$\sigma_\epsilon^2 = \frac{2B_L \mathcal{S}_N(0)}{\alpha^2} \text{sec}^2 \quad (28)$$

where  $B_L = \int_0^\infty |H(j2\pi f)|^2 df$  is the one-sided loop noise bandwidth. Hence, from Eqs. (17), (27), and (28)

$$\frac{\sigma_\epsilon^2}{T_H^2} = \frac{1}{3} - \frac{B_L W \left[1 - \frac{1}{3} \left(1 - \operatorname{erf}^2(\sqrt{R_H})\right)\right]}{R_H \operatorname{erf}^2(\sqrt{R_H})},$$

(fraction of a symbol)<sup>2</sup> (29)

Let

$$T = 2T_H \quad (30)$$

to relate the half-symbol to the full-symbol duration ( $T$ ).

Translating to the full-symbol period  $T$  yields the desired result for arbitrary window size.

$$\frac{\sigma_\epsilon^2}{T^2} = \frac{1}{6} \frac{W B_L \left[1 - \frac{1}{3} \left(1 - \operatorname{erf}^2\left(\sqrt{\frac{R}{2}}\right)\right)\right]}{R \operatorname{erf}^2\left(\sqrt{\frac{R}{2}}\right)},$$

(fraction of a symbol)<sup>2</sup> (31)

where  $R = PT/N_0 = E_S/N_0$ .

The value  $\sigma_\epsilon^2 / ((B_L T) T^2)$  is plotted in Fig. 3 for  $W = T/4$  versus  $R$ , the full-symbol SNR ( $E_S/N_0$ ), where  $E_S$  is the data symbol energy and  $N_0$  is the one-sided noise spectral density. The other loops will be discussed in the following sections. Taking the somewhat arbitrary window  $W = T/4$ , Eq. (31) becomes

$$\frac{\sigma_\epsilon^2}{T^2} = \frac{1}{24} \frac{B_L T \left[1 - \frac{1}{3} \left(1 - \operatorname{erf}^2\left(\sqrt{\frac{R}{2}}\right)\right)\right]}{R \operatorname{erf}^2\left(\sqrt{\frac{R}{2}}\right)} \text{(sec/sec)}^2 \quad (32)$$

The relationship  $W = T/4$  is used for all three loop window sizes, since at high data rates this would probably be the minimum size.

### III. Analysis of the Near-Optimum Mid-Transition-Tracking Manchester Symbol Synchronizer

This section describes a symbol synchronizer that is motivated by the optimum structure [4,5,6] for the mid-symbol transition and ignores the adjacent-symbol transition. Figure 4 illustrates this symbol synchronizer.

The received signal plus noise is modeled by Eq. (1) as

$$y(t) = \sqrt{P} \sum_{k=-\infty}^{\infty} b_k q(t - kT - \tau) + n(t) \quad (33)$$

with  $n(t)$  modeled as white Gaussian noise having spectral density  $N_0/2$ .

In order to analyze the symbol synchronizer depicted in Fig. 4, it is advantageous to segment the noise process into four contiguous regions as shown in Figure 5. The bottom portion of Fig. 5 illustrates the symbol sync reference mid-symbol transition point integration region (solid lines) and the actual mid-symbol transition point of the received signal (dashed lines). The pulse function  $q(t)$  is a complete Manchester symbol and  $b_k$  is the data sign ( $b_k = \pm 1$ ) from Eq. (1). The  $X$ -channel signal (see Fig. 4) can be written at the end of symbol  $b_k$ , as

$$X = \sqrt{P} \left[ b_k T - (b_{k-1} + 3b_k) \epsilon + (N_1 + N_2 - N_3 - N_4) \right], \quad \epsilon \geq 0 \quad (34a)$$

$$X = \sqrt{P} \left[ b_k T - (b_{k+1} + 3b_k) \varepsilon \right. \\ \left. + (N_1 + N_2 - N_3 - N_4) \right], \quad \varepsilon < 0 \quad (34b)$$

where it is assumed that  $|\varepsilon|$  ( $\varepsilon = \tau - \hat{\tau}$ ) is less than  $W$ . Additionally, it is assumed that  $\varepsilon$  is small in the computations that follow.

The  $Y$  channel produces the signal at the end of its integration time given by

$$Y = \left[ 2\sqrt{P} b_k \varepsilon + (N_2 - N_3) \right] \quad (35)$$

The product  $Z = XY = \hat{\varepsilon}$  is the estimate of the timing error over one pulse time and for small  $\varepsilon$  is given by

$$\hat{\varepsilon}(t) = 2PT\varepsilon g(t) + \sqrt{P} b_k T (N_2 + N_3) g(t) \\ + (N_1 + N_2 - N_3 - N_4) (N_2 + N_3) g(t) \quad (36)$$

where  $g(t)$  is the pulse function taking on the value of one at each loop update (see Fig. 8). This pulse function is constant over  $T$  sec and thus acts as a sample-and-hold function for the loop error signal.

Equation (36) gives the conditional mean value of  $\hat{\varepsilon}$  given  $\varepsilon$  over one update period as

$$E[\hat{\varepsilon}|\varepsilon] = 2PT\varepsilon \quad (37)$$

since

$$E[(N_1 + N_2 - N_3 - N_4)(N_2 + N_3)] = 0 \quad (38)$$

because

$$E[N_2^2] = E[N_3^2] \quad (39)$$

If the total noise term is denoted by  $N_T$ , where

$$N_T = \underbrace{\sqrt{PT}(N_2 + N_3)}_{N_A} + \underbrace{(N_1 + N_2 - N_3 - N_4)(N_2 + N_3)}_{N_B} \quad (40)$$

then it can be shown that the two noise components  $N_A$  and  $N_B$  are independent. That is,

$$E[N_A N_B] = \sqrt{PT} E[(N_2 + N_3) \\ \times (N_1 + N_2 - N_3 - N_4)(N_2 + N_3)] = 0 \quad (41)$$

since odd moments of zero mean Gaussian random variables are zero.

To obtain the loop equation, the estimate of the error is written as

$$\hat{\varepsilon} = 2PT\varepsilon + N_T g(t) \quad (42)$$

The loop timing estimate is given by

$$\hat{\tau} = \frac{KF(s)}{s} \hat{\varepsilon} = \frac{KF(s)}{s} [2PT\varepsilon + N_T g(t)] \quad (43)$$

where  $K$  is the loop gain of the symbol synchronizer and  $F(s)$  is the loop filter expressed as a Heaviside operator. Using

$$\tau - \hat{\tau} = \varepsilon \quad (44)$$

and for convenience  $\tau = 0$  in Eq. (43) yields

$$\varepsilon(t) = H(s) \left[ \frac{N_T g(t)}{2PT} \right] \quad (45)$$

where  $H(s)$  is the closed-loop transfer function and

$$H(s) = \frac{KF(s)/s}{1 + KF(s)/s} \quad (46)$$

When the loop noise bandwidth  $B_L$  is small compared to the symbol rate, the linearized tracking error can be approximated by

$$\sigma_{\varepsilon}^2 = \frac{2B_L \mathcal{S}_{N_T}(0)}{4P^2 T^2} \quad (47)$$

where  $\mathcal{S}_{N_T}(0)$  is the spectral density at  $f = 0$  of the cyclostationary process  $N_T(t)$ , with  $N_T(t)$  defined below.

Thus, it is necessary to evaluate the spectral density of the noise process at  $f = 0$ . The noise process can be written as

$$N_T(t) = \sum_{k=-\infty}^{\infty} N_T(k)g(t - kT) \quad (48)$$

and thus the  $N_T(t)$  process is cyclostationary. It can be made stationary by averaging over time. Thus, the auto-correlation function of the stationary equivalent process is given by

$$R(\zeta) = \frac{1}{T} \int_0^T E[N_T(t)N_T(t + \zeta)] dt \quad (49)$$

Thus

$$R_{N_T}(\zeta) = \left. \begin{aligned} &= \sigma^2_{N_T} \left(1 - \frac{|\zeta|}{T}\right) && \text{for } |\zeta| \leq T \\ &= 0 && \text{otherwise} \end{aligned} \right\} \quad (50)$$

Since

$$\mathcal{P}_{N_T}(0) = \int_{-\infty}^{\infty} R(\zeta) d\zeta = T\sigma^2_{N_T} \quad (51)$$

one has for Eq. (47) that

$$\sigma^2_{\epsilon} = \frac{B_L T \sigma^2_{N_T}}{2P^2 T^2} \quad (52)$$

Evaluating  $\sigma^2_{N_T}$  yields

$$\sigma^2_{N_T} = \frac{N_0 W P T^2}{2} + \frac{N_0^2 W T}{4} \quad (53)$$

Therefore, using Eq. (53) in Eq. (52), with  $R = PT/N_0$ , the normalized tracking error is given by

$$\frac{\sigma^2_{\epsilon}}{T^2} = \frac{B_L W}{4R} \left[1 + \frac{1}{2R}\right] (\text{fraction of a chip})^2 \quad (54)$$

The normalized tracking error is plotted in Fig. 3. In the comparison,  $W$  was set equal to  $T/4$  to be consistent

with the other two loops considered here. Thus, Eq. (54) becomes

$$\frac{\sigma^2_{\epsilon}}{T^2} = \frac{B_L T}{16R} \left[1 + \frac{1}{2R}\right] \quad (55)$$

## IV. Analysis of the Mid-Symbol- and Adjacent-Symbol-Tracking Manchester Symbol Synchronizer

Figure 6 shows the symbol synchronizer discussed in this section. The functions  $h(t)$  and  $1 - h(t)$  are shown in Fig. 7. The upper branch performs the same function as the previous section and operates on the mid-symbol transitions. However, in addition, the lower section operates on the adjacent-symbol transition points. The lower and upper sections are used to update the loop at twice the symbol rate, unlike the loop discussed in the previous section, which is updated every symbol time.

### A. Mid-Symbol Error Detection

First consider the upper two branches of the mid-symbol error detector illustrated in Fig. 6. A symbol and loop timing diagram is illustrated in Fig. 7. Initially,  $W_B$  and  $W_M$  denote the between-symbol and mid-symbol windows, respectively. Again the signal is modeled as described in Eq. (1).

The signal denoted  $X$  in Fig. 6 is given (for small  $\epsilon$ ) by

$$X = \sqrt{P}(b_k T - (b_{k-1} + 3b_k)|\epsilon|) + (N_{00} + N_1 + N_2 - N_3 - N_4 - N_5) \quad (56)$$

where the integration regions are as indicated in Fig. 8. The  $Y$ -channel output (for small  $\epsilon \geq 0$ ) is given by

$$Y = \left(2\sqrt{P}b_k \epsilon + N_2 + N_3\right) \quad (57)$$

where

$$N_{00} = \int_0^{W_B/2} n(t) dt \quad (58)$$

with the time origin taken at the point where the synchronizer loop starts the  $k$ th symbol for simplicity of notation. In addition, there are the definitions

$$N_1 = \int_{W_B/2}^{T/2 - W_M/2} n(t) dt \quad (59)$$

$$N_2 = \int_{T/2 - W_M/2}^{T/2} n(t) dt \quad (60)$$

$$N_3 = \int_{T/2}^{T/2 + W_M/2} n(t) dt \quad (61)$$

$$N_4 = \int_{T/2 + W_M/2}^{T - W_B/2} n(t) dt \quad (62)$$

$$N_5 = \int_{T - W_B/2}^T n(t) dt \quad (63)$$

$$N_6 = \int_T^{T + W_B/2} n(t) dt \quad (64)$$

Thus, the first error signal  $Z_1$  is given by

$$Z_1 = XY = \hat{\varepsilon}_1$$

and from Eqs. (56) and (57), for small  $|\varepsilon|$

$$\hat{\varepsilon}_1 = 2PT\varepsilon + \overbrace{\sqrt{P}b_k T(N_2 + N_3)}^{N_A} + \overbrace{(N_{00} + N_1 + N_2 - N_3 - N_4 - N_5)(N_2 + N_3)}^{N_B} \quad (65)$$

over  $T/2$  seconds. Thus, showing the explicit time dependence of the error signal with time

$$\hat{\varepsilon}_1(t) = \sum_{k=-\infty}^{\infty} (2PT\varepsilon(k) + N_A(k) + N_B(k))h(t - kT) \quad (66)$$

where  $h(t)$  is defined to be unity in the region  $t \in (0, T/2)$  and zero in the region  $t \in (T/2, T)$ . The functions  $h(t)$  and  $1 - h(t)$  are plotted in Fig. 8 along with  $p(t)$ ,  $g(t)$ , and  $q(t)$ . The quantities  $\varepsilon(k)$ ,  $N_A(k)$ , and  $N_B(k)$  are the values of the respective variables at the  $k$ th symbol time.

Equation (66) is the contribution of the mid-symbol error detector composed of the upper two branches in Fig. 6. Furthermore, from Eq. (65) the mean value of  $\hat{\varepsilon}_1$  over one symbol time is given by

$$E[\hat{\varepsilon}_1|\varepsilon] = 2PT\varepsilon \quad (67)$$

## B. Adjacent-Symbol Error Detection

Now consider the adjacent-symbol transition detector depicted in the lower half of Fig. 6:

$$U'_k = 2b_k\sqrt{P}(\tau - \hat{\tau}) + N_5 + N_6 \quad (68)$$

For the upper branch of the adjacent-symbol detector, the detected half-symbols are given by

$$\hat{a}_{2K+1} = \text{sgn}[-b_k(T/2 - 2|\varepsilon|) + N_3 + N_4 + N_5] \quad (69)$$

and

$$\hat{a}_{2k+2} = \text{sgn}[b_{k+1}(T/2 - 2|\varepsilon|) + N_6 + N_7] \quad (70)$$

In Fig. 6 a scale factor of  $\beta\sqrt{PT}$  has been included in the upper branch of the lower half of the figure. Its purpose is to make the units the same ( $\sqrt{PT}$ ) for the upper and lower halves, with a scale factor of  $\beta$  ( $0 \leq \beta$ ) used to adjust the relative proportion of each error signal.

Hence, the upper branch of the between-half-symbol transition detector denoted by  $I_k$  can be expressed in terms of the  $a_k$  sequence as

$$I_k = \left( \frac{-\hat{a}_{2k+1} + \hat{a}_{2k+2}}{2} \right) \beta\sqrt{PT} \quad (71)$$

The estimate of the error signal from the lower half is therefore given by

$$\hat{\varepsilon} = U_k I_k \quad (72)$$

or

$$\begin{aligned} \hat{\varepsilon}_2 = & \left( 2b_k \sqrt{P} (\tau - \hat{\tau}) + N_5 + N_6 \right) \\ & \times \left( \frac{-\hat{a}_{2k+1} + \hat{a}_{2k+2}}{2} \right) \beta \sqrt{PT} \end{aligned} \quad (73)$$

To obtain the conditional mean value of  $\hat{\varepsilon}_2$  conditioned on  $\varepsilon$ , a simplifying assumption is made. First, neglect the small correlation of  $(N_5 + N_6)$  with  $\hat{a}_{2k+1}$  and  $\hat{a}_{2k+2}$ . Second, assume that  $\varepsilon$  is very small in magnitude. Thus, letting  $\varepsilon = \tau - \hat{\tau}$  obtains

$$E[\hat{\varepsilon}_2|\varepsilon] = 2\beta PT \left( \frac{1}{2} \right) \left\{ (1 - PE_H)^2 - PE_H^2 \right\} \quad (74)$$

or

$$E[\hat{\varepsilon}_2|\varepsilon] = \beta PT \varepsilon (1 - 2PE_H) \quad (75)$$

where the symbol error rate  $PE_H$  is the same as Eq. (16), so that

$$1 - 2PE_H = \operatorname{erf} \left( \sqrt{R/2} \right) \quad (76)$$

where it was assumed that  $\varepsilon = 0$  in the expression for  $PE_H$ . Thus,

$$E[\hat{\varepsilon}_2|\varepsilon] = \beta PT \varepsilon \operatorname{erf} \left( \sqrt{R/2} \right) \quad (77)$$

### C. Tracking Performance of the Combined Loop Signal

The total error signal that drives the loop filter  $F(s)$  of Fig. 5 is given by

$$\begin{aligned} \hat{\varepsilon}(t) = & \frac{1}{2} \left[ \beta PT \operatorname{erf} \left( \sqrt{R/2} \right) + 2PT\varepsilon \right] \\ & + \sum_{k=-\infty}^{\infty} \{ N_1(kT)h(t - kT) \\ & + N_2(kT)[1 - h(t - kT)] \} \end{aligned} \quad (78)$$

where the first term in Eq. (78) is the mean value of  $\varepsilon(t)$ , and where the second term is the noise process with  $\varepsilon$  assumed to be zero. The noise terms are given by

$$\begin{aligned} N_1(kT) = & \sqrt{P} b_k T (N_2 + N_3) \\ & + (N_{00} + N_1 + N_2 - N_3 - N_4 - N_5)(N_2 + N_3) \end{aligned} \quad (79)$$

and

$$N_2(kT) = \beta \sqrt{PT} \left( \frac{-\hat{a}_{2k+1} + \hat{a}_{2k+2}}{2} \right) (N_5 + N_6) \quad (80)$$

The symbol synchronizer forms the estimate  $\hat{\tau}$  of the received signal delay and can be expressed by

$$\begin{aligned} \hat{\tau} = & \frac{KF(s)}{s} \hat{\varepsilon} \\ = & \frac{KF(s)}{s} \left\{ PT + \frac{1}{2} \beta PT \operatorname{erf} \left( \sqrt{\frac{R}{2}} \right) \right\} \varepsilon + \frac{KF(s)}{s} \{ N(t) \} \end{aligned} \quad (81)$$

where

$$N(t) = \sum_{k=-\infty}^{\infty} \{ N_1(kT)h(t - kT) + N_2(kT)[1 - h(t - kT)] \} \quad (82)$$

where  $K$  is the loop gain including the phase detector gain, and the ratio  $F(s)/s$  is the loop filter expressed in the Heaviside polynomial divided by the filtering effect of the VCO ( $1/s$ ). Noting that  $\varepsilon = \tau - \hat{\tau}$  and assuming that  $\tau = 0$  for convenience leads to

$$\varepsilon(t) = H(s) \left[ \frac{N(t)}{PT + \frac{\beta}{2} PT \operatorname{erf} \left( \sqrt{\frac{R}{2}} \right)} \right] \quad (83)$$

where  $H(s)$  is the closed-loop transfer function of the symbol synchronizer loop. Following the usual practice, it is assumed that the one-sided loop noise bandwidth  $B_L$  is much smaller than the symbol rate, so that the variance of the linearized tracking error in Eq. (83) can be determined from the expression

$$\sigma_\varepsilon^2 = \frac{2B_L \mathcal{S}_N(0)}{\left[ PT + \frac{\beta}{2} PT \operatorname{erf} \left( \sqrt{\frac{R}{2}} \right) \right]^2} \quad (84)$$

where  $\mathcal{S}_N(0)$  is the spectral density of the noise process  $N(t)$ . To evaluate the spectral density, the autocorrelation function of the noise is determined. Since the noise process is cyclostationary, time is averaged over one period to obtain a stationary process.

$$\begin{aligned}
R_N(\tau) &\equiv \frac{1}{T} \int_0^T R_N(t + \tau, t) dt \\
&= E \int_0^T \left\{ \sum_{k=-\infty}^{\infty} \{N_1(kT)h(t + \tau - kT) \right. \\
&\quad \left. + N_2(kT)[1 - h(t + \tau - kT)]\} \right. \\
&\quad \times \sum_{\ell=-\infty}^{\infty} \{N_1(\ell T)h(t - \ell T) \\
&\quad \left. + N_2(\ell T)[1 - h(t - \ell T)]\} \right\} dt \quad (85)
\end{aligned}$$

assuming for convenience that  $E[N_1(kT)N_2(\ell T)] \cong 0$  for all  $k$  and  $\ell$  (even though there is a small correlation between them). Letting  $\ell = k + m$  obtains

$$\begin{aligned}
R_N(\tau) &= \frac{1}{T} \int_0^T \left\{ \sum_{m=-\infty}^{\infty} R_{N_1}(mT) \right. \\
&\quad \times \sum_{k=-\infty}^{\infty} h(t + \tau - kT)h(t - (k + m)T) \left. \right\} dt \\
&+ \frac{1}{T} \int_0^T \left\{ \sum_{m=-\infty}^{\infty} R_{N_2}(mT) \sum_{k=-\infty}^{\infty} [1 - h(t + \tau - kT)] \right. \\
&\quad \left. \times [1 - h(t - (k + m)T)] \right\} dt \quad (86)
\end{aligned}$$

Additionally,  $R_{N_1}(mT) = 0$  for all  $m \neq 0$  since  $N_1(kT)$  and  $N_1((k + 1)T)$  are based on integrations over disjoint time intervals. Furthermore,  $R_{N_2}(mT) \cong 0$  for all  $m \neq 0$  since  $N_2(kT)$  and  $N_2((k + 1)T)$  only have  $W_B/2 \ll T$  sec

in common over the adjacent (full) symbol times. Using these two conditions obtains

$$\begin{aligned}
R_N(\tau) &= \sigma_{N_1}^2 \frac{1}{T} \int_0^T h(t + \tau)h(t)dt + \sigma_{N_2}^2 \\
&\quad \times \frac{1}{T} \int_0^T [1 - h(t + 1)][1 - h(t)] dt \quad (87)
\end{aligned}$$

Completing the averaging,

$$R_N(t) = \frac{\sigma_{N_1}^2}{T} R(\tau) + \frac{\sigma_{N_2}^2}{T} R(\tau) \quad (88)$$

where

$$R(\tau) = \int_0^T h(t + \tau)h(t) dt \quad (89)$$

and is illustrated in Fig. 9. Therefore, the spectral density of the noise process is given by

$$\mathcal{S}_N(f) = \frac{\sigma_{N_1}^2}{T} |S(f)|^2 + \frac{\sigma_{N_2}^2}{T} |S(f)|^2 \quad (90)$$

where

$$S(f) = \int_0^T e^{-i\omega\tau} d\tau = \frac{T}{2} e^{-i\omega T/4} \frac{\sin(\pi f T/2)}{(\pi f T/2)} \quad (91)$$

so that

$$|S(f)|^2 = \frac{T^2}{4} \frac{\sin^2(\pi f T/2)}{(\pi f T/2)^2} \quad (92)$$

From Eqs. (88) and (90),  $\mathcal{S}_N(0)$  is given by

$$\mathcal{S}_N(0) = \left( \frac{\sigma_{N_1}^2 + \sigma_{N_2}^2}{4} \right) T \quad (93)$$

Thus, to evaluate the tracking-error variance it is necessary to evaluate  $\sigma_{N_1}^2$  and  $\sigma_{N_2}^2$ . First,  $\sigma_{N_1}^2$  is determined. From Eq. (79),

$$\begin{aligned}
N_1 &= \sqrt{P}b_k T(N_2 + N_3) \\
&+ (N_{00} + N_1 + N_2 - N_3 - N_4 - N_5)(N_2 + N_3)
\end{aligned} \tag{94}$$

Since the two terms are uncorrelated and have zero mean values, the variance of  $N_1$  is given by the sum of the variances in Eq. (94). If the first grouping of Eq. (94) is denoted as  $N_A$  and the second grouping as  $N_B$ , then  $E[N_A^2]$  and  $E[N_B^2]$  can be evaluated. Consider the former:

$$E[N_A^2] = PT^2 E[(N_2 + N_3)^2] = PT^2 \frac{N_0}{2} W_M \tag{95}$$

Now consider the latter term with

$$\begin{aligned}
E[N_B^2] &= E[N_2^2 N_{00}^2] + E[N_2^2 N_1^2] + E[N_2^4] + E[N_2^2 N_3^2] \\
&+ E[N_2^2 N_4^2] + E[N_2^2 N_5^2] - 4E[N_2^2 N_3^2] \\
&+ E[N_3^2 N_{00}^2] + E[N_3^2 N_1^2] + E[N_3^2 N_2^2] \\
&+ E[N_3^4] + E[N_3^2 N_4^2] + E[N_3^2 N_5^2]
\end{aligned} \tag{96}$$

After simplifying, one obtains

$$E[N_B^2] = \frac{N_0^2 W_M T}{4} \tag{97}$$

and

$$\sigma_{N_1}^2 = PT^2 \frac{N_0}{2} W_M + \frac{N_0^2 W_M T}{4} \tag{98}$$

Now consider the computation of  $\sigma_{N_2}^2$ . From Eq. (80),

$$N_2 = \beta \sqrt{P} T \left[ \frac{-\hat{a}_{2k+1} + \hat{a}_{2k+2}}{2} \right] (N_5 + N_6) \tag{99}$$

To evaluate  $N_2$ , the small correlations between  $N_5$  and  $\hat{a}_{2k+1}$  and  $N_6$  and  $\hat{a}_{2k+2}$  are neglected so that

$$\sigma_{N_2}^2 = \beta^2 PT^2 E[N_5^2 + N_6^2] E \left[ \left( \frac{-\hat{a}_{2k+1} + \hat{a}_{2k+2}}{2} \right)^2 \right] \tag{100}$$

or

$$\sigma_{N_2}^2 = \frac{N_0}{4} W_B \beta^2 PT^2 \tag{101}$$

since the transition detector term has an average value of 1/2. So,

$$\frac{\sigma_{\epsilon}^2}{T^2} = \frac{\beta_L T \left[ 1 + 2R \left( 1 + \frac{\beta^2}{2} \right) \right]}{32R^2 \left[ 1 + \frac{\beta}{2} \operatorname{erf} \left( \sqrt{\frac{R}{2}} \right) \right]^2} (\text{fraction of a symbol})^2 \tag{102}$$

Notice that when  $\beta = 0$ , this result is the same as Eq. (55), as it should be! Since  $\beta$  is a parameter, it can be varied to minimize Eq. (102). Figure 3 illustrates the results for this symbol sync loop plotted versus  $R$  in decibels. For this loop,  $\beta\sqrt{PT}$  must be known a priori to obtain optimum performance. However, the parameter  $\beta$  is not very sensitive. For example, at  $R = -12$  dB,  $\beta_{\text{opt}} = 1.75$  yields a normalized tracking error of  $7.52 \text{ sec}^2/\text{sec}^2$ , and at the value  $\beta = 1$ , the normalized tracking error becomes  $7.73 \text{ sec}^2/\text{sec}^2$ . However, at  $\beta = 0$  (mid-transition detector only), the normalized tracking error becomes 8.84, the same as the second symbol sync loop considered. In fact, when  $\beta = 0$ , the two curves are identical as noted above.

Therefore, since  $T$  would be known precisely a priori, and since  $P$  would be known to within 10 to 20 percent, it seems that setting  $\beta = 1$  would allow very close to optimum ( $\beta_{\text{opt}}$ ) performance. Furthermore,  $\beta_{\text{opt}}$  is equal to approximately 1 at  $E_S/N_0 \geq 7$  dB, so that under most reasonable conditions of the Block V receiver setting  $\beta = 1$  is optimum.

## V. Conclusion

All three symbol-synchronization tracking loops offer fairly similar performance. The hybrid loop called optimum Manchester is better (low tracking error) for  $R \leq$

0 dB than the other two loops. However, for  $R \geq 0$ , the NRZ $\times 2$  loop and the hybrid optimum Manchester loop are essentially equal in performance. For the hybrid optimum Manchester loop to work, the power of the signal has to be estimated to provide the weighting  $\beta\sqrt{PT}$  in Fig. 6 with  $\beta$  set equal to unity.

Although the hybrid Manchester loop is optimum, it is not clear that the extra hardware requirement of this loop is warranted. It is necessary to compare the actual estimated tracking losses for each loop based on the requirements to determine if the complexity of the hybrid loop is justified and if it is best at high data rates.

## References

- [1] J. K. Holmes, *Coherent Spread Spectrum Systems*, New York: Wiley-Interscience, 1982.
- [2] W. J. Hurd and T. O. Anderson, "Digital Transition Tracking Symbol Synchronizer for Low SNR Coded Systems," *IEEE Trans. Comm. Tech.*, vol. COM-18, no. 4, pp. 141-147, April 1970.
- [3] L. Franks, *Signal Theory*, New York: Prentice-Hall Inc., 1969.
- [4] A. Cellier, "An All-Digital Manchester Symbol Synchronizer for Low SNR," *NAECON 75 Record, Proceedings of the National Aerospace and Electronics Conference*, Dayton, Ohio, pp. 403-409, June 10-12, 1975.
- [5] M. K. Simon and W. C. Lindsey, "Tracking Performance of Symbol Synchronizers for Manchester Coded Data," *IEEE Trans. on Comm.*, vol. 25, no. 4, pp. 398-408, April 1977.
- [6] B. H. Batson, H. Vang, A. Cellier, and W. L. Lindsey, "An All-Digital Manchester Symbol Synchronizer for Space Shuttle," *National Telecommunications Conference*, San Diego, California, pp. 724-733, December 2-4, 1974.

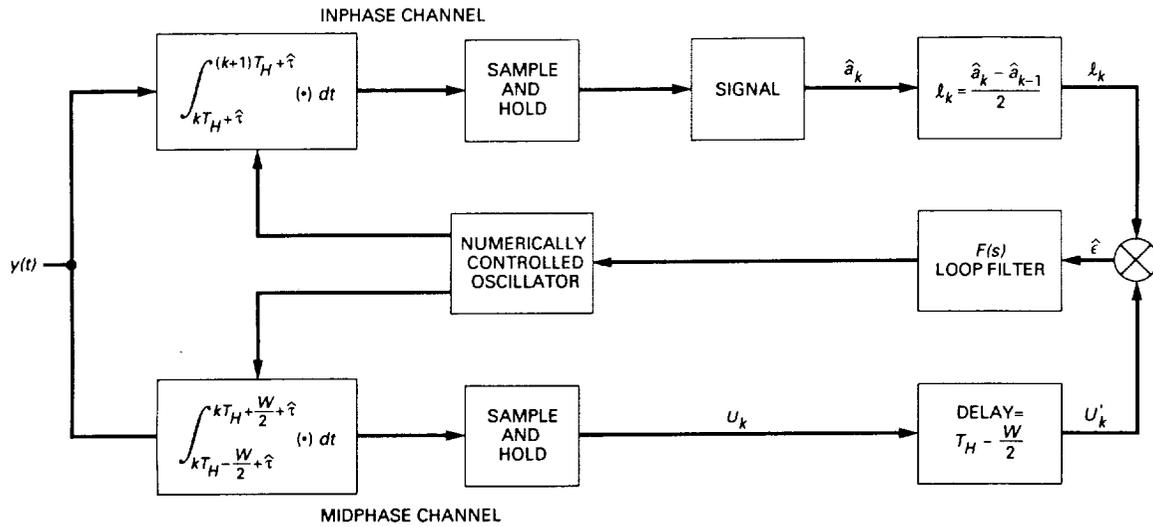
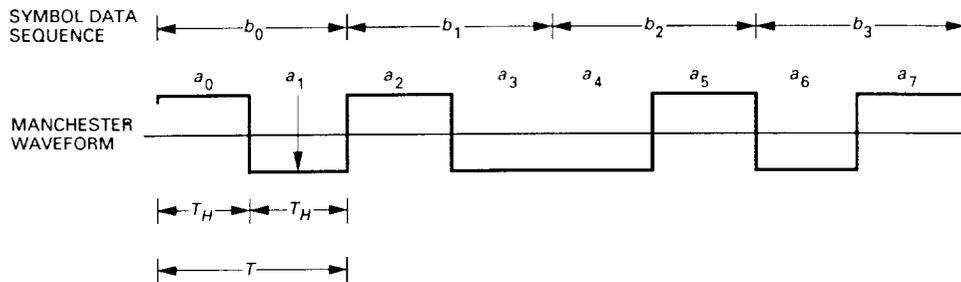


Fig. 1. The NRZ  $\times 2$  Manchester symbol synchronizer.



SYMBOL DATA SEQUENCE:  $b_0 = 1, b_1 = 1, b_2 = -1, b_3 = -1, \dots$   
 PULSE SEQUENCE:  $a_0 = 1, a_1 = -1, a_2 = 1, a_3 = -1, a_4 = -1, a_5 = 1, a_6 = -1, a_7 = 1, \dots$

Fig. 2. The relationship between the full-symbol data sequence and the Manchester half-symbols.

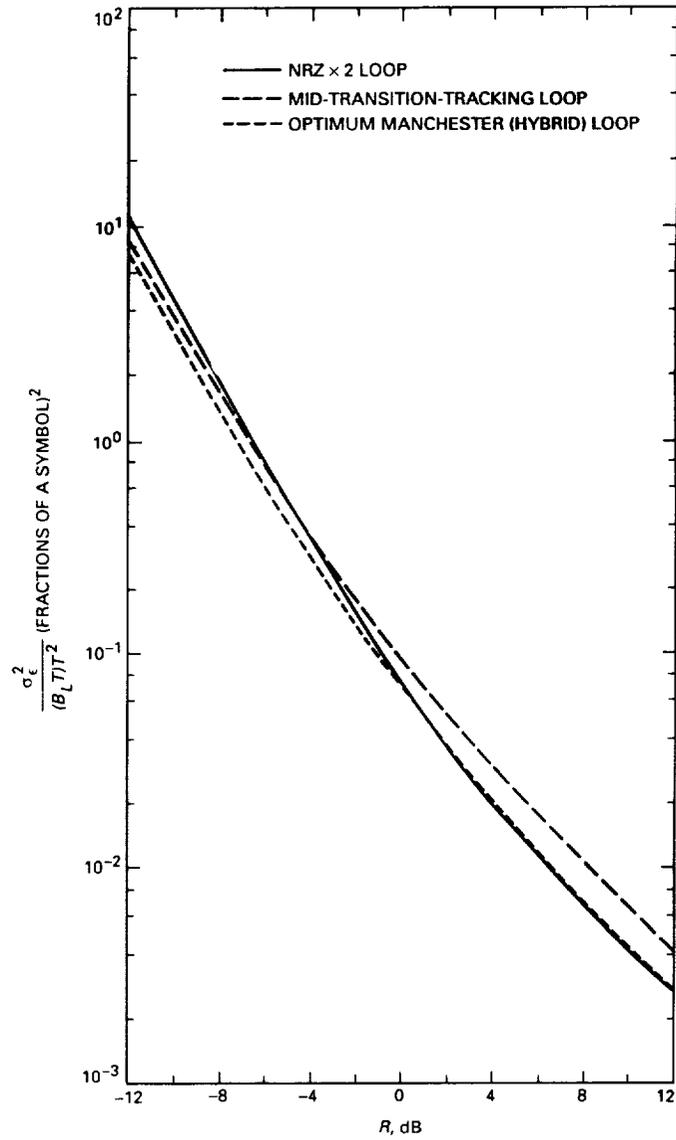


Fig. 3. A comparison of the three Manchester symbol-synchronization loops.

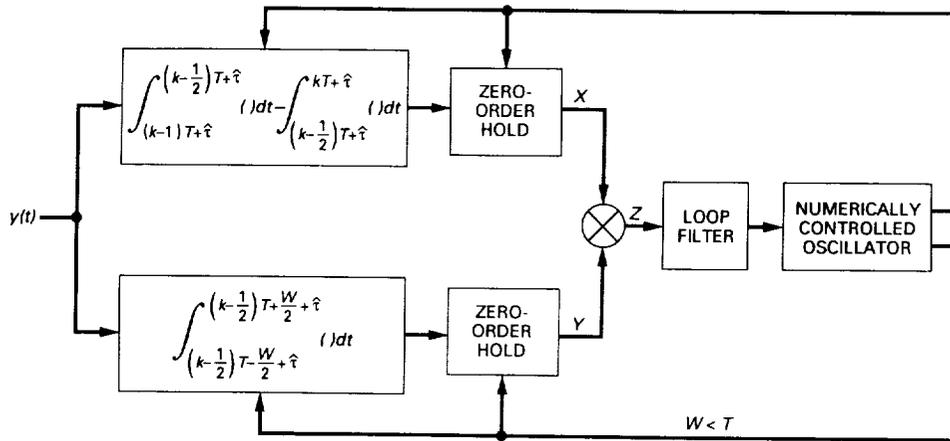


Fig. 4. The mid-transition-tracking Manchester symbol synchronizer.

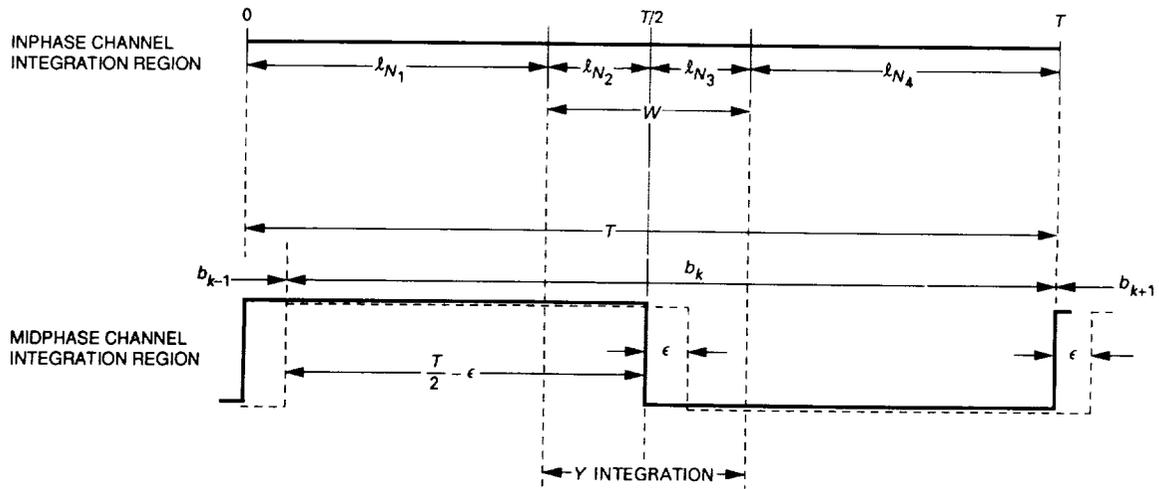


Fig. 5. The inphase channel and the midphase channel integration region for the Manchester symbol synchronizer (shown for  $W/2 > \epsilon > 0$ ).

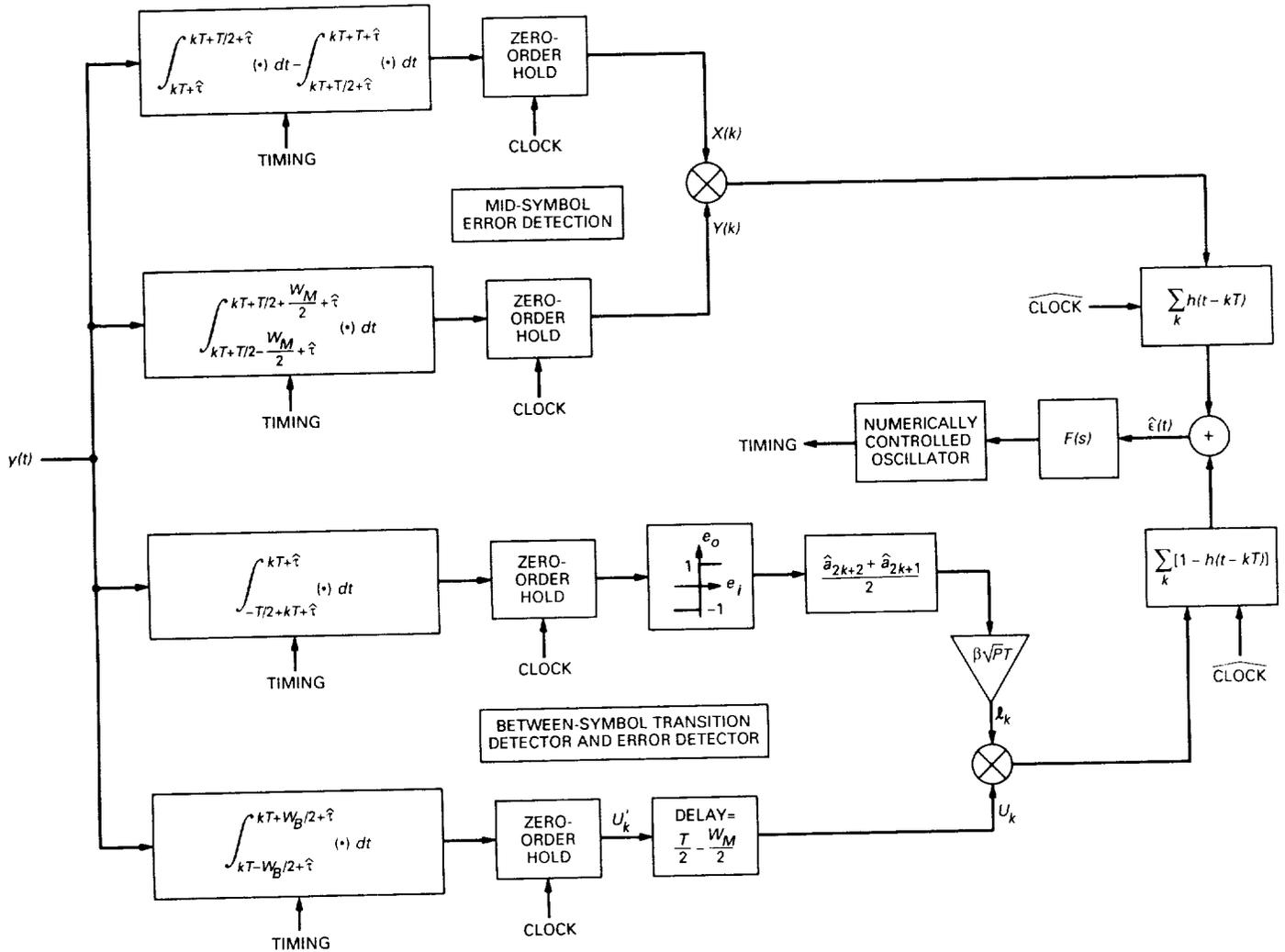


Fig. 6. A Manchester symbol synchronizer for tracking mid-symbol and adjacent-symbol transitions.

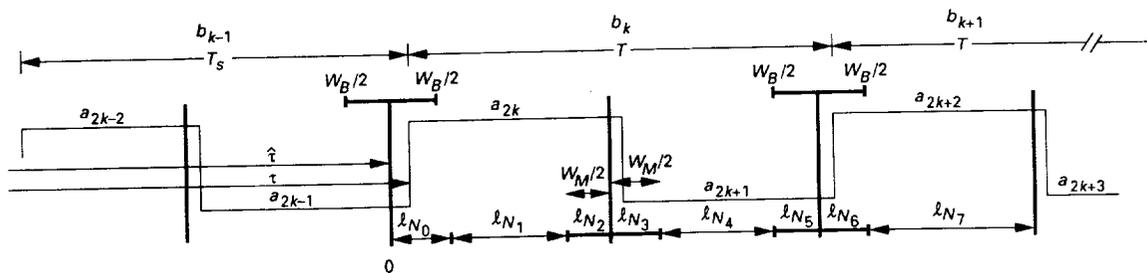
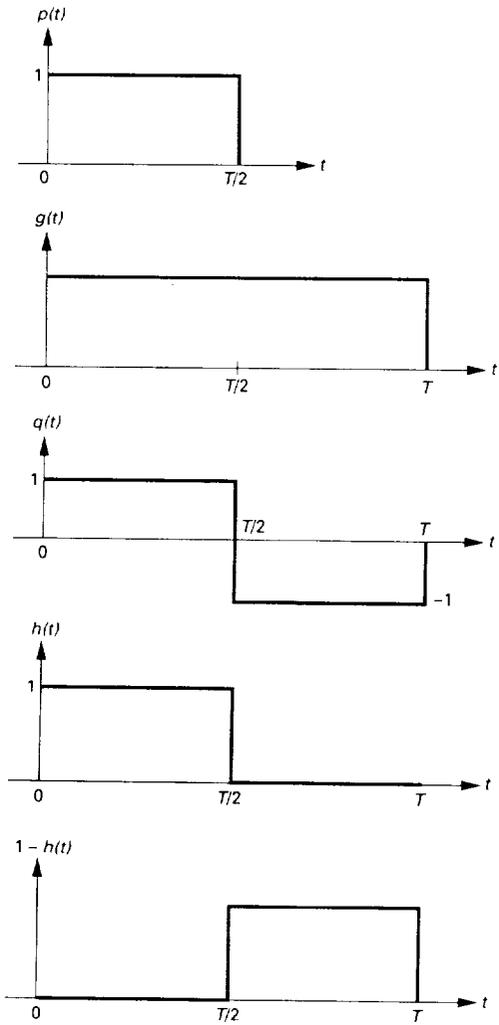
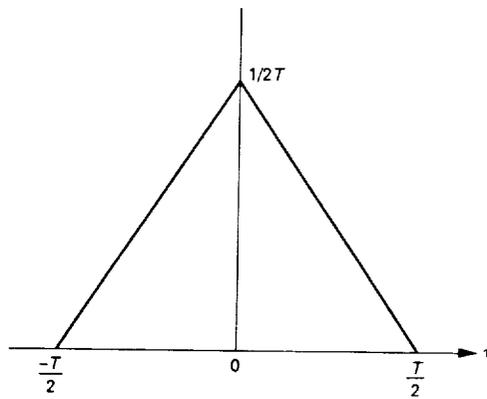


Fig. 7. Symbol and loop timing diagram. Loop timing is indicated in dark lines and received signal is indicated in light lines.



**Fig. 8. Repetitive pulses used in the noise analysis.**



**Fig. 9. Autocorrelation function for the  $h(t)$  multiplexing symbols.**

516-32  
33224 P-18

N91-18324

3322451

# Open-Loop Frequency Acquisition for Suppressed-Carrier Biphase Signals Using One-Pole Arm Filters

B. Shah and J. K. Holmes  
Radio Frequency and Microwave Subsystems Section

*This article discusses open-loop frequency-acquisition performance for suppressed-carrier binary phase shift keyed signals in terms of the probability of detecting the carrier frequency offset when the arms of the Costas loop detector have one-pole filters. The approach, which does not require symbol timing, uses fast Fourier transforms (FFTs) to detect the carrier frequency offset. The detection probability, which depends on both the 3-dB arm filter bandwidth and the received symbol signal-to-noise ratio, is derived and is shown to be independent of symbol timing. It is shown that the performance of this technique is slightly better than other open-loop acquisition techniques which use integrators in the arms and whose detection performance varies with symbol timing.*

## I. Introduction

Acquiring and tracking binary phase shift keyed (BPSK) signals in the absence of a residual carrier is one of the many functional requirements that must be met by the Block V receiver [1], the Deep Space Network's (DSN's) next-generation receiver presently under development. The question of how to reduce the frequency error  $\Delta f$  between the received and predicted carrier frequency to within the pull-in range of the Block V receiver's digital Costas loop (the loop used to demodulate BPSK signals) was initially addressed in [2]. The open-loop acquisition techniques discussed in [2] estimate  $\Delta f$  by performing fast Fourier transforms (FFTs) on the phase detector output of Costas-type loops. Because the techniques in [2] use integrate-and-dump arm filters, their performance in terms of probability of detecting  $\Delta f$  depends strongly on symbol timing errors, that is, the receiver's estimate of where a symbol epoch starts and ends. In particular, the performance of these techniques, which worsens with

increasing symbol timing errors, motivates evaluating system performance when the integrate-and-dump arm filters are replaced by one-pole arm filters that are independent of symbol timing. This article considers only the estimation of the initial frequency error and does not treat the acquisition of the phase and frequency after the frequency error has been removed and the loop is closed (see Fig. 1). Frequency and phase acquisition when the frequency error is less than one-half the closed-loop noise bandwidth should occur within a few inverse closed-loop noise bandwidths [3,4].

Figure 1 is a functional block diagram of the open-loop frequency-acquisition technique under consideration. The arm filters are assumed to be one-pole filters. The error signal  $z'(t)$  is sampled every  $T$  sec ( $T$  sec is the duration of a symbol), accumulated, and fast-Fourier-transformed to obtain the error signal spectrum. The probability of detecting a tone in white Gaussian noise is well known [5]

and can easily be computed as a function of the error-signal SNR. The results of [5] can be applied in a straightforward manner because, as shown in Appendix A, the error sequence  $z(n)$  out of the integrate-and-dump filter is composed of a sinusoid at  $2\Delta f$  Hz plus noise which is approximately white and approximately Gaussian. The absolute frequency difference  $|2\Delta f|$  can be estimated (it is assumed in the analysis to follow that  $\Delta f T \ll 1$ ) after detecting the error signal in the spectrum at the output of the Costas loop phase detector. Only the absolute value of  $2\Delta f$  can be estimated because taking the product of the  $I$  and  $Q$  arms doubles the frequency error and removes its sign. The sign ambiguity problem can be resolved by offsetting the frequency so that the error  $\Delta f$  always has a known sign.

This article only considers the case of detecting and subsequently estimating the frequency error when the error signal is restricted to a single FFT bin. It also assumes that the error signal is always present. Consequently, the article does not address the question of how to choose a threshold, which is the power level that the FFT bin with the maximum power must exceed before a signal can be declared to be present.

When a tone is known to be present and restricted to a single bin, the maximum likelihood (ML) estimate of the tone is to observe the FFT magnitude spectrum and to select the frequency bin with the maximum power [5]. Note that, in practice, it is unlikely that the error signal will be restricted to a single FFT bin.

In a typical deep-space operation, the Doppler shift is accompanied by a Doppler rate which is removed by ramping the local oscillator phase based on Doppler rate predicts. If the Doppler rate cannot be predicted precisely, the error signal will drift in frequency during the FFT observation time, causing it to smear over multiple bins. Clearly, the extent of the smearing depends on the size of the Doppler rate error. On the other hand, if the Doppler rate is known very accurately so that the signal does not drift appreciably during the observation interval, the signal may be present in multiple bins due to not sampling the discrete-time Fourier transform at its peak (see Chapter 8 of [6]). Restricting the error signal to a single bin greatly simplifies the analysis. Furthermore, it is appropriate because the intent is not to quantify absolute performance but rather to compare performance between the lowpass technique and the integration techniques of [2].

Although other models are not precluded, sampling of  $z'(t)$  is modeled as a  $T$ -sec integrate-and-dump filter, because doing so defines the signal-to-noise ratio (SNR) of

the error sequence  $z(n)$  in the same bandwidth as the alternative acquisition techniques studied in [2]. Defining SNRs in the same bandwidth is important when comparing performance between different schemes. Since the arm filters do not require symbol timing, and taking the product of the  $I$  and  $Q$  arms prior to the integrate-and-dump filter gives rise to a signal component (a tone at  $2\Delta f$ ) whose amplitude does not depend on symbol timing, the performance of the one-pole technique is independent of symbol timing errors. For convenience, the integrator in Fig. 1 is shown as synchronous with symbol transitions although the analysis does not depend on this fact. In fact, the value of integration time  $T'$  does not have to be equal to  $T$  as long as  $\Delta f T' \ll 1$ . However, as shown in Section II, performance depends directly on the normalized bandwidth  $R$ , defined as the ratio of the 3-dB arm filter bandwidth to the data rate and the symbol SNR.

A mathematical model of the one-pole technique is developed in Section II. Its performance in terms of probability of detecting the frequency difference between the received and predicted carrier frequency is derived in Section III and discussed in Section IV. Conclusions are stated in Section V.

## II. Mathematical Model

The received suppressed-carrier BPSK signal, downconverted to an appropriate intermediate frequency (IF), can be modeled as

$$r(t) = \sqrt{2P}d(t) \sin(\omega_i t + \theta_i) + n(t) \quad (1)$$

where  $P$  is the received power in watts,  $\omega_i$  is the IF radian frequency expressed in rad/sec,  $\theta_i$  is the signal phase in rad, and  $d(t)$  is the transmitted data stream given by

$$d(t) = \sum_{k=-\infty}^{+\infty} d_k p(t - kT) \quad (2)$$

where  $p(t)$  is the baseband non-return-to-zero (NRZ) pulse limited to  $T$  sec and  $d_k$  represents the equally likely  $\pm 1$  binary symbols. The narrow-band noise process  $n(t)$  can be expressed as

$$n(t) = \sqrt{2}n_c(t) \cos(\omega_i t + \theta_i) - \sqrt{2}n_s(t) \sin(\omega_i t + \theta_i) \quad (3)$$

where  $n_c(t)$  and  $n_s(t)$  are statistically independent stationary band-limited white Gaussian noise processes with one-sided spectral density  $N_0$  watts/Hz and one-sided bandwidth  $W$  Hz. The signal  $r(t)$  is demodulated by in-phase and quadrature references,  $\sqrt{2} \sin(\omega_o t + \theta_o)$  and

$\sqrt{2}\cos(\omega_o t + \theta_o)$ , tuned to the predicted carrier frequency and then filtered by a lowpass filter. Neglecting higher frequency terms, the demodulated signals  $i(t)$  and  $q(t)$  in Fig. 1 can be represented as (see Chapter 5 of [7] for an example):

$$i(t) = \text{Re} \left[ \sqrt{P}d(t)e^{j\phi}e^{j\Delta\omega t} \right] - \text{Re} \left[ n_c(t)\frac{1}{j}e^{j\phi}e^{j\Delta\omega t} \right] - \text{Re} \left[ n_s(t)e^{j\phi}e^{j\Delta\omega t} \right] \quad (4)$$

and

$$q(t) = \text{Re} \left[ \sqrt{P}d(t)\frac{1}{j}e^{j\phi}e^{j\Delta\omega t} \right] + \text{Re} \left[ n_c(t)e^{j\phi}e^{j\Delta\omega t} \right] - \text{Re} \left[ n_s(t)\frac{1}{j}e^{j\phi}e^{j\Delta\omega t} \right] \quad (5)$$

where  $\text{Re}$  is the real part of a complex number,  $\Delta\omega = 2\pi\Delta f$  ( $\Delta f \triangleq f_i - f_o$ ) is the radian frequency error to be estimated, and  $\phi \triangleq \theta_i - \theta_o$  is the phase error. Let  $D(\omega)$ ,  $N_s(\omega)$ , and  $N_c(\omega)$  be the Fourier transforms of the baseband signals  $d(t)$ ,  $n_s(t)$ , and  $n_c(t)$ . Then the Fourier representations of  $i(t)$  and  $q(t)$  are given by

$$I(\omega) = \frac{\sqrt{P}}{2} \left[ D(\omega - \Delta\omega)e^{j\phi} + D^*(-\omega - \Delta\omega)e^{-j\phi} \right] - \frac{1}{2j} \left[ N_c(\omega - \Delta\omega)e^{j\phi} - N_c^*(-\omega - \Delta\omega)e^{-j\phi} \right] - \frac{1}{2} \left[ N_s(\omega - \Delta\omega)e^{j\phi} + N_s^*(-\omega - \Delta\omega)e^{-j\phi} \right] \quad (6)$$

and

$$Q(\omega) = \frac{\sqrt{P}}{2j} \left[ D(\omega - \Delta\omega)e^{j\phi} - D^*(-\omega - \Delta\omega)e^{-j\phi} \right] + \frac{1}{2} \left[ N_c(\omega - \Delta\omega)e^{j\phi} + N_c^*(-\omega - \Delta\omega)e^{-j\phi} \right] - \frac{1}{2j} \left[ N_s(\omega - \Delta\omega)e^{j\phi} - N_s^*(-\omega - \Delta\omega)e^{-j\phi} \right] \quad (7)$$

where  $X^*$  is the conjugate of the complex number  $X$ . Let  $H(\omega)$  be the transfer function of the arm filters in Fig. 1.

Then, the filter outputs in the inphase and quadrature arms are given as

$$Z_i(\omega) = \sqrt{\frac{P}{2}} \left[ \tilde{D}(\omega - \Delta\omega)e^{j\phi} + \tilde{D}^*(-\omega - \Delta\omega)e^{-j\phi} \right] - \frac{1}{2j} \left[ \tilde{N}_c(\omega - \Delta\omega)e^{j\phi} - \tilde{N}_c^*(-\omega - \Delta\omega)e^{-j\phi} \right] - \frac{1}{2} \left[ \tilde{N}_s(\omega - \Delta\omega)e^{j\phi} + \tilde{N}_s^*(-\omega - \Delta\omega)e^{-j\phi} \right] \quad (8)$$

and

$$Z_q(\omega) = \frac{\sqrt{P}}{2j} \left[ \tilde{D}(\omega - \Delta\omega)e^{j\phi} - \tilde{D}^*(-\omega - \Delta\omega)e^{-j\phi} \right] + \frac{1}{2} \left[ \tilde{N}_c(\omega - \Delta\omega)e^{j\phi} + \tilde{N}_c^*(-\omega - \Delta\omega)e^{-j\phi} \right] - \frac{1}{2j} \left[ \tilde{N}_s(\omega - \Delta\omega)e^{j\phi} - \tilde{N}_s^*(-\omega - \Delta\omega)e^{-j\phi} \right] \quad (9)$$

where  $\tilde{D}(\omega)$ ,  $\tilde{N}_c(\omega)$ , and  $\tilde{N}_s(\omega)$  are defined, as in [7], as

$$\tilde{D}(\omega) = D(\omega)H(\omega + \Delta\omega) \quad (10)$$

$$\tilde{N}_c(\omega) = N_c(\omega)H(\omega) \quad (11)$$

$$\tilde{N}_s(\omega) = N_s(\omega)H(\omega) \quad (12)$$

Equations (8) and (9) represent the output of the arm filters when the signal portion of the input has bandpass spectra centered at  $\Delta\omega$ , and the input noise is bandlimited white noise with bandwidth  $W \gg B + \Delta f$  where  $B$  Hz is the 3-dB bandwidth of  $H(\omega)$ . Equations (8) and (9) are applicable since the demodulated signals  $i(t)$  and  $q(t)$  have signal spectra centered at  $\Delta\omega$  during acquisition. The time-domain representations of  $Z_i(\omega)$  and  $Z_q(\omega)$  are given by

$$z_i(t) = \text{Re} \left[ \sqrt{P}\tilde{d}(t)e^{j\phi}e^{j\Delta\omega t} \right] + \tilde{n}_i(t) \quad (13)$$

and

$$z_q(t) = \text{Re} \left[ \sqrt{P}\tilde{d}(t)\frac{1}{j}e^{j\phi}e^{j\Delta\omega t} \right] + \tilde{n}_q(t) \quad (14)$$

where

$$\begin{aligned} \tilde{n}_i(t) = & -Re \left[ \tilde{n}_c(t) \frac{1}{j} e^{j\phi} e^{j\Delta\omega t} \right] \\ & - Re \left[ \tilde{n}_s(t) e^{j\phi} e^{j\Delta\omega t} \right] \end{aligned} \quad (15)$$

$$\begin{aligned} \tilde{n}_q(t) = & Re \left[ \tilde{n}_c(t) e^{j\phi} e^{j\Delta\omega t} \right] \\ & - Re \left[ \tilde{n}_s(t) \frac{1}{j} e^{j\phi} e^{j\Delta\omega t} \right] \end{aligned} \quad (16)$$

and where  $\tilde{d}(t)$ ,  $\tilde{n}_s(t)$ , and  $\tilde{n}_c(t)$  are the inverse Fourier transforms of  $\tilde{D}(\omega)$ ,  $\tilde{N}_s(t)$ , and  $\tilde{N}_c(t)$ . It can be shown from Eq. (10) that when  $\Delta\omega \neq 0$ ,  $\tilde{d}(t)$  is in general complex; when  $\Delta\omega = 0$ , it is real.

### III. One-Pole Arm Filters

Assuming that the arm filters in Fig. 1 are one-pole filters with 3-dB bandwidth  $B$  Hz, the transfer function  $H(\omega)$  is given by

$$H(\omega) = \frac{\omega_B}{j\omega + \omega_B} \quad (17)$$

where  $\omega_B = 2\pi B$  is the 3-dB filter bandwidth in radians. For such a filter, the error signal  $z'(t) = z_i(t)z_q(t)$  can be computed as (see Appendix A):

$$z'(t) = A(P, \Delta\omega T, R) \sin(2\Delta\omega t + 2\phi + \xi) + n'_{eff}(t) \quad (18)$$

The amplitude  $A$  given by Eq. (A-5) is a function of many variables, including the normalized bandwidth  $R = BT$ , which is the ratio of the 3-dB bandwidth of the one-pole filter to the data rate  $R_s = 1/T$ . It can be shown that  $A$  does not change significantly with  $\Delta fT$  when  $\Delta fT \ll 1$ . Since this article only considers the case  $\Delta fT \ll 1$ , the dependence of  $A$  on  $\Delta\omega T$  is omitted in subsequent expressions to allow a simpler notation. Since the interest here is the magnitude of the error signal, the phase  $\xi$  in Eq. (18) is not relevant and consequently is not included here. The effective noise  $n'_{eff}(t)$  into the integrator is defined as

$$n'_{eff}(t) \triangleq n'_{ss}(t) + n'_{sn}(t) + n'_{nn}(t) \quad (19)$$

where the self-noise due to the signal times the signal product,  $n'_{ss}(t)$ , is given by Eq. (A-7); the noise due to the signal times the noise product,  $n'_{sn}(t)$ , is given by Eq. (A-8);

and the noise due to the noise times the noise product,  $n'_{nn}(t)$ , is given by Eq. (A-9). The signal-signal noise,  $n'_{ss}(t)$ , which is a consequence of intersymbol interference (ISI), has two terms. The first term has a continuous spectrum and zero mean, whereas the second term gives rise to line spectra at harmonics (not including the fundamental harmonic) of the symbol rate. The noise  $n'_{sn}(t)$  is the low-pass filtered signal response at time  $t$  (due to the present as well as previously transmitted pulses) times the thermal noise filtered by a lowpass filter. Lastly, the noise-noise process  $n'_{nn}(t)$  is the product of the filtered thermal noise in the inphase arm and the quadrature-phase arm.

These noises are independent of each other since data and noise are assumed to be independent. Consequently, the average power of the effective noise is the sum of the average power of each of the noise processes above. Expressions for average noise power are given by evaluating the autocorrelation functions of Eqs. (A-11) through (A-13) at  $\tau = 0$ .

Referring to Fig. 1, the process  $z'(t)$  is integrated and dumped to obtain the sequence  $z(n)$ . Assuming that the frequency error is much smaller than the data rate, that is,  $\Delta fT \ll 1$ , it is shown in Eq. (A-15) that the error sequence  $z(n)$  can be represented as

$$z(n) = A(P, R) \sin[2\Delta\omega(nT + T/2) + 2\phi + \xi] + n_{eff}(n) \quad (20)$$

where  $A(P, R)$  is as defined in Eq. (A-5) and the effective noise sequence  $n_{eff}(n)$  is defined to be

$$n_{eff}(n) \triangleq n_{ss}(n) + n_{sn}(n) + n_{nn}(n) \quad (21)$$

where the noises  $n_{ss}(n)$ ,  $n_{sn}(n)$ , and  $n_{nn}(n)$  are given by Eqs. (A-16) through (A-18). The discrete autocorrelation functions of these noises, assuming  $\Delta\omega = 0$ , are derived in Appendix C and listed in Eqs. (A-19) through (A-21). These functions are exact except for the autocorrelation of  $n_{ss}(n)$ , which neglects the negligible second term in Eq. (A-7). Since data and noise are independent, the effective noise power is given as  $\sigma_{eff}^2$ , where

$$\sigma_{eff}^2 = \sigma_{ss}^2 + \sigma_{sn}^2 + \sigma_{nn}^2 \quad (22)$$

and where Eqs. (A-19) through (A-21) for  $R_{ss}(m)$ ,  $R_{sn}(m)$ , and  $R_{nn}(m)$  yield

$$\sigma_{ss}^2 = \frac{P^2}{8\pi^2 R^2} \left( -1 + \frac{e^{-2\pi R} + e^{2\pi R}}{2} \right)^2 \left( \frac{e^{-4\pi R}}{1 - e^{-4\pi R}} \right) \quad (23)$$

$$\sigma_{sn}^2 = P\sigma^2 \left[ 1 - \frac{1}{4\pi R} \left( 5 - \frac{9}{4\pi R} \right) + \frac{e^{-2\pi R}}{4} \left( 1 + \frac{1}{2\pi R} - \frac{17}{8\pi^2 R^2} \right) - \frac{1}{16\pi^2 R^2} \left( e^{-4\pi R} + \frac{e^{-6\pi R}}{2} \right) \right] \quad (24)$$

$$\sigma_{nn}^2 = \sigma^4 \left[ \frac{\pi R}{2} - \frac{1}{8} (1 - e^{-4\pi R}) \right] \quad (25)$$

where  $\sigma^2 = N_0/2T$ . The signal-to-noise ratio of the error sequence  $z(n)$ , defined as signal power divided by the noise power, is given by

$$SNR_z(R) = \frac{A^2(P, R)}{2\sigma_{eff}^2} \quad (26)$$

where, after one uses Eq. (A-5) for  $A(P, R)$  and Eq. (22) for  $\sigma_{eff}$ , the term  $SNR_z$  can be written in terms of the received symbol SNR as

$$SNR_z(R) = \frac{|C_0(R)|^2}{8C_{ss}(R) + \frac{4}{SNR}C_{sn}(R) + \frac{2}{SNR^2}C_{nn}(R)} \quad (27)$$

where  $SNR \triangleq PT/N_0$ ,  $|C_0(R)|^2$  is given by Eq. (A-6), and where  $C_{ss}(R) = \sigma_{ss}^2/P^2$ ,  $C_{sn}(R) = \sigma_{sn}^2/P\sigma^2$ , and  $C_{nn}(R) = \sigma_{nn}^2/\sigma^4$ . The quantities  $C_{ss}(R)$ ,  $C_{sn}(R)$ , and  $C_{nn}(R)$  depend on  $R$  because  $\sigma_{ss}^2$ ,  $\sigma_{sn}^2$ , and  $\sigma_{nn}^2$  are functions of  $R$ .

As expected, for a given SNR,  $SNR_z$  is primarily a function of  $R$ . Equation (27) is exact except for  $C_{ss}(R)$ , which is approximate because  $\sigma_{ss}^2$  is an approximation: see derivation of Eq. (A-19). Although it is not proven, the  $\sigma_{ss}^2$  given by Eq. (23) is believed to be a slight upper bound to the self-noise power for NRZ pulses. At symbol SNRs below 0 dB,  $SNR_z$  is a very accurate expression for the true error sequence SNR. This is because in this region the sum of  $C_{sn}(R)$  and  $C_{nn}(R)$ , which are scaled by the inverse of SNR and are exact, dominates  $C_{ss}(R)$ . For SNRs above 0 dB,  $SNR_z$  is believed to be a slight lower bound to the true error sequence SNR, because in this case  $C_{ss}(R)$  dominates the other two terms in the denominator of Eq. (27).

The objective of the analysis presented here is to enable the designer to choose an arm filter bandwidth  $R$  for this scheme that optimizes the probability of detecting a tone in white Gaussian noise, that is, detecting the frequency error between the incoming signal and the local oscillator frequency so that the reference frequency can be moved to the input frequency. Then the loop can be closed to speed acquisition. This is accomplished by choosing an  $R$  that optimizes the SNR of a tone in white Gaussian noise. The error sequence SNR given by Eq. (27) represents the SNR of a tone imbedded in noise  $n_{eff}(n)$ , which is neither exactly white nor exactly Gaussian but can be assumed to be both in practice. This is because the random variable in each FFT bin, which results from summing  $N$  appropriately weighted random variables at the FFT input, tends toward a Gaussian random variable for large values of  $N$ . The correlation coefficient  $\rho$  for a one-symbol separation is less than 0.2 for symbol SNRs below 0 dB and  $R > 0.3$ . When  $\rho \leq 0.2$ , the assumption of independent samples out of the integrate-and-dump is valid. Also, it is true that  $n_{eff}(n)$  is essentially white when  $R > 0.5$  for SNRs above 0 dB. Thus, for any optimum  $R$  it is essentially true that the integrate-and-dump output sequence is white and so the results of [5] apply for the probability of correct detection of the frequency error.

## IV. Numerical Results and Discussion

### A. SNR Degradation

Figure 2 depicts SNR degradation  $D$  versus normalized bandwidth  $R$ . Degradation is defined as the reduced error signal SNR given by Eq. (26) relative to the SNR of the error signal of an "ideal" Costas loop. An ideal Costas loop has integrate-and-dump arm filters with (magically) perfect symbol timing instead of lowpass filters. In [2] it is shown that the error signal SNR for an ideal Costas loop detector  $SNR_i$  is given by

$$SNR_i = \frac{SNR}{4 + 2/SNR} \quad (28)$$

where  $SNR \triangleq PT/N_0$  is the received symbol signal-to-noise ratio. In mathematical terms, the degradation is given as

$$D(R) = \frac{SNR_i}{SNR_z(R)} \quad (29)$$

Note that degradation defined in this way is a number greater than one, which indicates an actual loss. Figures 2(a) through 2(e) show degradation  $D(R)$  versus a relative 3-dB bandwidth  $R$  for various values of SNR. As

expected, there is an optimum value for  $R$  (i.e., arm filter bandwidth) that minimizes degradation and thus maximizes error signal SNR. Values of  $R$  greater or less than this optimum value decrease error signal SNR because they increase the noise power faster than the error signal power. From Fig. 2 it is clear that lower symbol SNRs result in smaller optimum  $R$  values.

The reference SNR,  $SNR_i$  given by Eq. (28), is plotted versus symbol SNR in Fig. 3. Figures 3 and 2 relate received symbol SNR to the error signal SNR,  $SNR_z$ .

## B. Probability of Detection

The outlier probability (missed detection probability)  $q$ , defined as the probability that the magnitude of any FFT noise-only bin exceeds the magnitude of the signal-plus-noise bin, is given by  $q = 1 - p$ , where  $p$ , the probability of detecting the correct frequency offset, is given by ( $M = N/2$ ):

$$p = \int_0^\infty 2M(SNR_z)y e^{-M(SNR_i)(y^2+1)} \times I_0[2My(SNR_z)] \left[1 - y^2 e^{-M(SNR_i)}\right]^M dy \quad (30)$$

where  $N$  is the FFT size (and  $M$  is one-half the FFT size) and  $I_0()$  is the modified Bessel function of the first kind. Thus, this probability for  $p$  is given by [5] except that the factor of  $M - 1$  is replaced by  $M$  in the last bracketed term. If  $M$  were  $M - 1$ , the expressions in [5] and Eq. (30) would be identical. The results in [5] are slightly different from Eq. (30) because the results in [5] are derived for the more general complex FFT case. Figure 4 depicts  $q$  versus SNR where SNR corresponds to the signal-to-noise ratio at the FFT input (i.e.,  $SNR_z$ ). Figure 4 applies only when the noise prior to the FFT operation is white and Gaussian. In the case described here, the error signal noise component  $n_{eff}(n)$  is essentially white and Gaussian. Since the  $N$ -point FFT has input samples at the symbol rate, the frequency bin size is

$$\Delta f_{BIN} = \frac{1}{NT} \quad (31)$$

with  $(1/T)$  as the integrate-and-dump rate which is assumed to be the symbol rate in this analysis in order to compare the results with [2]. Thus, when the correct frequency bin is detected, the actual frequency error, assuming that the error signal is at the center of a bin, is reduced

to a maximum of  $\Delta f_{BIN}/2$ . As long as the maximum frequency error is less than one-half of the loop bandwidth (that is,  $N$  is large enough), and

$$\frac{\Delta f_{BIN}}{2} \leq \frac{B_L}{2} \quad (32)$$

or equivalently

$$\Delta f_{BIN} \leq B_L \quad (33)$$

where  $B_L$  is the one-sided (closed-loop) bandwidth of the Costas loop, the frequency and phase pull-in should be on the order of a few inverse  $B_L$ 's.

The following example illustrates how to use the curves presented in this section to compute the probability of correctly detecting the frequency error "seen" by the loop. Suppose that the received symbol SNR is 0 dB. Then, from Fig. 3, the error signal SNR of an ideal Costas loop is -11 dB. Assuming that  $R = 0.3$ , from Fig. 2(c),  $D(0.3) = 0.3$  dB and the error signal SNR  $SNR_z = -11.3$  dB. Finally, from Fig. 4, the probability of incorrectly detecting the actual frequency error for a 1024-point FFT is  $1.3 \times 10^{-6}$ .

## V. Conclusion

This article has described a method of determining the probability of correctly identifying the frequency error between the incoming suppressed-carrier signal and the frequency of the Costas loop oscillator in order to aid in the frequency acquisition of the suppressed-carrier signal. The detector chosen for estimating the frequency error is the error detector of the Costas loop, which is used for tracking the suppressed-carrier signal. The error signal is not fed back to the loop filter and numerically controlled oscillator (NCO), but rather sent to an  $N$ -point FFT. The FFT then estimates the frequency error, the Costas loop is closed, and the loop NCO is adjusted in frequency to reduce the initial frequency error to a very small value.

Knowing the symbol SNR allows the determination of the  $SNR_i$  and, from Figs. 2(a) through 2(e), the additional degradation  $D$ . The addition of these two terms yields  $SNR_z$ , which is the abscissa entry on the plot of the probability of incorrectly detecting the actual frequency error (Fig. 4). One minus this probability yields the probability of correctly detecting the initial frequency error between the received signal and the rest frequency of the Costas loop NCO.

This scheme compares favorably to the three methods suggested in [2]. In particular, at SNRs that are  $\leq 0$  dB, the staggered approach of [2] is best (minimum degradation). Comparing the average loss (averaged over timing error) at SNRs  $\leq 0$  dB, the one-pole arm filter approach is about 0.5 dB better. At SNRs  $\geq 0$  dB it is better by more than 0.5 dB.

Whether the one-pole filter approach or one of the methods suggested in [2] will be used for the Block V receiver depends on hardware considerations, since the performance of the best integrate-and-dump arm filter technique is nearly comparable to the one-pole arm filter approach and is easily switched to the optimum tracking configuration with integrate-and-dump arm filters.

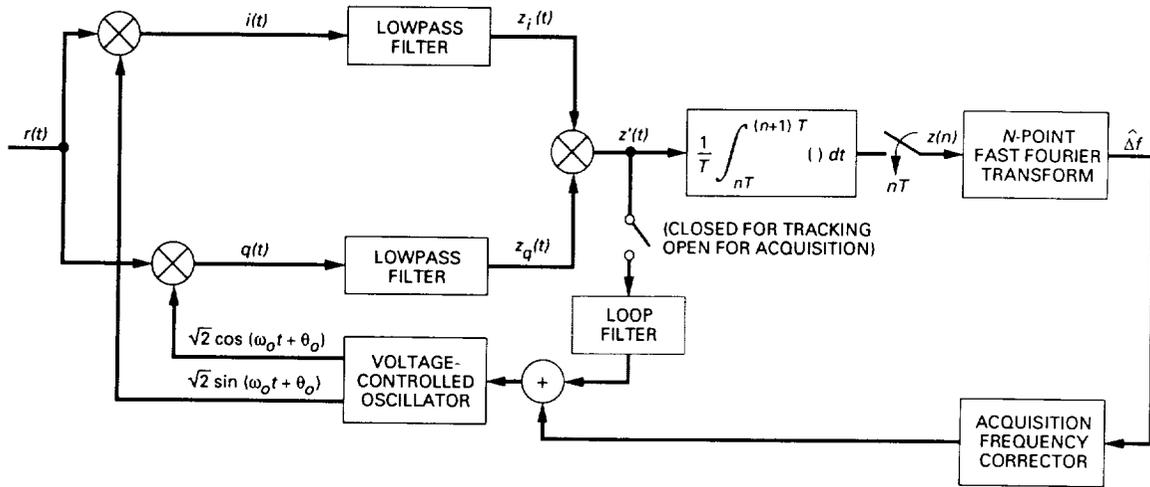


Fig. 1. Costas loop error detector to determine and correct the frequency error between the input and the reference oscillators.

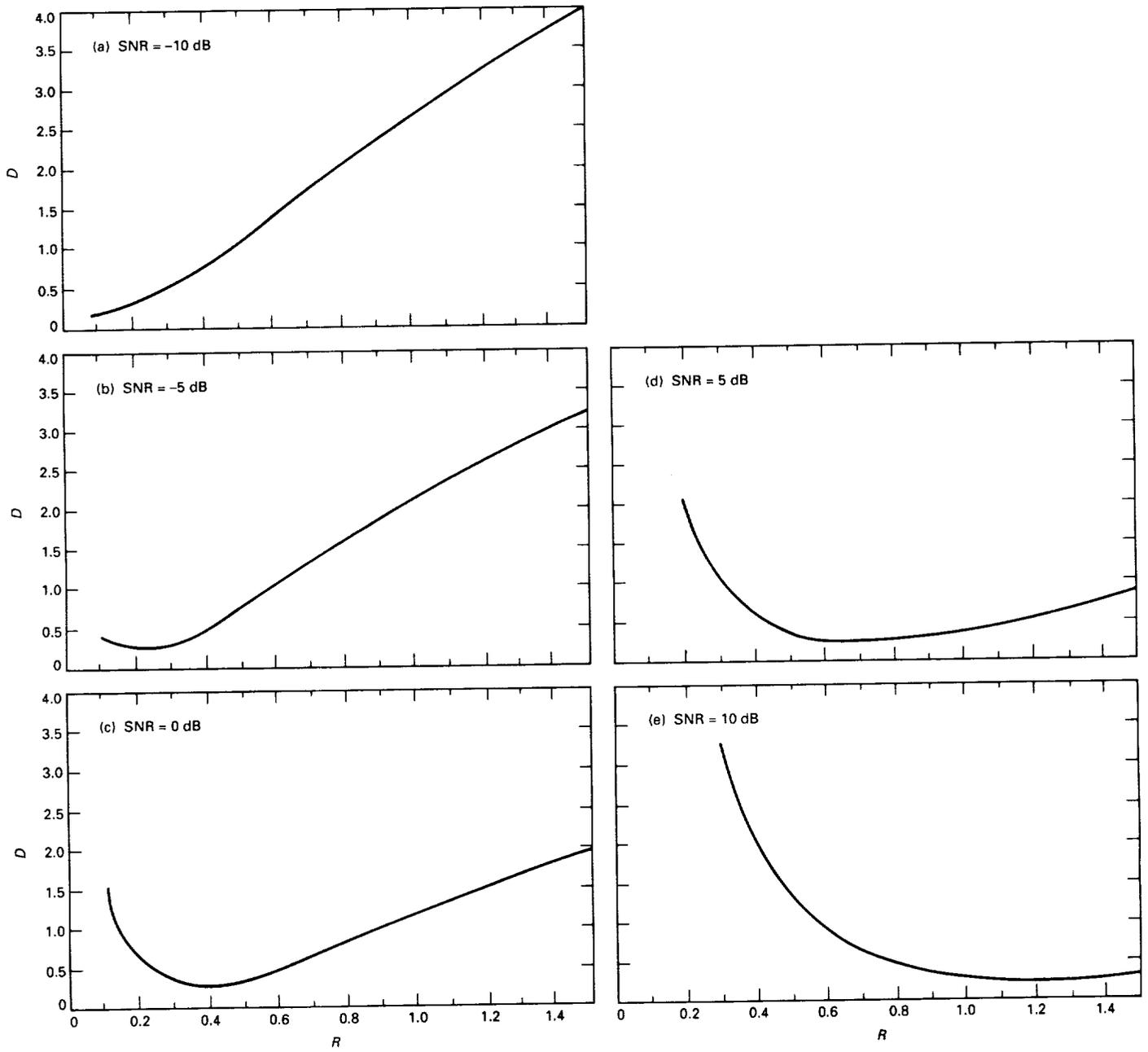


Fig. 2. Signal-to-noise degradation  $D$  versus normalized bandwidth  $R$  for  $0 \leq \Delta\omega T \leq 0.1$  and for various values of SNR.

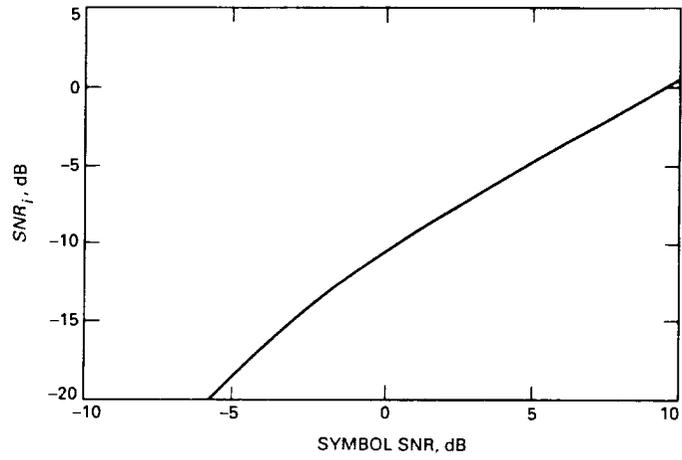


Fig. 3. Signal-to-noise ratio of the "ideal" integrate-and-dump Costas detector versus symbol SNR.

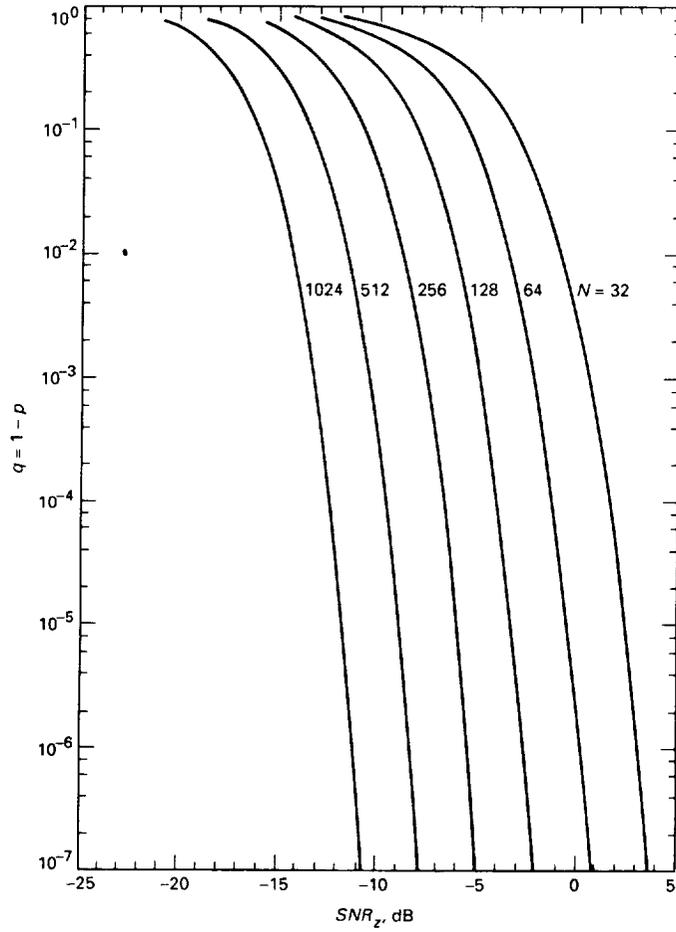


Fig. 4. Probability of incorrectly detecting the frequency error with a Costas loop error detector.

## Appendix A

### Derivation of the Error Sequence $z(n)$

Figure 1 shows how lowpass filtered signals are collapsed into an error signal that can be Fourier transformed to estimate the frequency error  $\Delta f$ . The received signal  $r(t)$  downconverted to an appropriate IF is given by Eqs. (1)-(3). The signal  $r(t)$  is first demodulated and then lowpass filtered to produce the functions  $z_i(t)$  and  $z_q(t)$ , which are given by Eqs. (13) and (14). Referring to Fig. 1, the error signal  $z'(t)$  is given by

$$\begin{aligned} z'(t) &= z_i(t)z_q(t) \\ &= \text{Im} \left[ \frac{P}{2} \tilde{d}^2(t) e^{j2(\Delta\omega t + \phi)} \right] + \text{Re} \left[ \sqrt{P} \tilde{d}(t) e^{j(\Delta\omega t + \phi)} \right] \tilde{n}_q(t) + \text{Re} \left[ \sqrt{P} \tilde{d}(t) \frac{1}{j} e^{j(\Delta\omega t + \phi)} \right] \tilde{n}_i(t) + \tilde{n}_i(t) \tilde{n}_q(t) \end{aligned} \quad (\text{A-1})$$

where  $\text{Re}$  and  $\text{Im}$  are, respectively, the real and imaginary parts of a complex number, and where  $\tilde{n}_i(t)$  and  $\tilde{n}_q(t)$  are given by Eqs. (15) and (16). The quantity  $\tilde{d}^2(t)$  is given as

$$\tilde{d}^2(t) = \left| \tilde{d}^2(t) \right| e^{j\xi(t)} \quad (\text{A-2})$$

It can be shown from Eq. (10) that  $\tilde{d}^2(t)$ , which is generally a complex quantity, reduces to a real quantity when  $\Delta\omega = 0$ . In this case, Eq. (A-1) reduces to  $(P\tilde{d}^2(t)/2) \sin 2\phi$  plus noise. This is the well-known form for the error signal of a Costas loop [8], where the quantity  $\tilde{d}(t)$  is the filtered version of the baseband data  $d(t)$ . When  $\Delta\omega \neq 0$ , the function  $\tilde{d}^2(t)$  for a one-pole filter is derived in Appendix B to be

$$\tilde{d}^2(t) = \sum_{n=-\infty}^{\infty} C_n(R) e^{j(2n\pi t/T)} + \sum_{k=-\infty}^{\infty} \sum_{\substack{l=-\infty \\ l \neq k}}^{\infty} d_k d_l \tilde{p}(t-lT) \tilde{p}(t-kT) \quad (\text{A-3})$$

where  $\tilde{p}(t)$  and  $C_n(R)$  are respectively given by Eqs. (B-8) and (B-9). Using Eq. (A-3) for  $\tilde{d}^2(t)$  and Eq. (B-4) for  $\tilde{d}(t)$  in Eq. (A-1), the error signal can be written as

$$z'(t) = A(P, R) \sin(2\Delta\omega t + 2\phi + \xi) + n'_{eff}(t) \quad (\text{A-4})$$

where

$$A(P, R) = \frac{P}{2} |C_0(R)| \quad (\text{A-5})$$

and where, from Eq. (B-10), the squared magnitude of  $C_0(R)$  can be computed to be

$$\begin{aligned} |C_0(R)|^2 &= \left[ \frac{1}{1 + \left(\frac{\Delta\omega T}{2\pi R}\right)^2} \right]^2 \left\{ 1 + \frac{1}{1 + \left(\frac{\Delta\omega T}{2\pi R}\right)^2} \left[ \frac{1}{(2\pi R)^2} (1 + e^{-4\pi R} - 2e^{-2\pi R} \cos(\Delta\omega T)) \right. \right. \\ &\quad \left. \left. - \frac{2}{2\pi R} \left( 1 - e^{-2\pi R} \left( \cos(\Delta\omega T) - \frac{\Delta\omega T}{2\pi R} \sin(\Delta\omega T) \right) \right) \right] \right\} \end{aligned} \quad (\text{A-6})$$

Since the interest here is in the spectrum of the error signal, the phase  $\xi$  in Eq. (A-4) is not relevant. Consequently, it is not included here.

The effective noise  $n'_{eff}(t)$  in Eq. (A-4) is defined in Eq. (19), where

$$n'_{ss}(t) = Im \left[ \frac{P}{2} \sum_{k=-\infty}^{\infty} \sum_{\substack{l=-\infty \\ l \neq k}}^{\infty} d_k d_l \tilde{p}(t - kT) \tilde{p}(t - lT) e^{j2(\Delta\omega t + \phi)} \right] + Im \left[ \frac{P}{2} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} C_n(R) e^{j(2n\pi t/T)} e^{j2(\Delta\omega t + \phi)} \right] \quad (A-7)$$

$$n'_{sn}(t) = Re \left[ \sqrt{P} \sum_{k=-\infty}^{\infty} d_k \tilde{p}(t - kT) e^{j(\Delta\omega t + \phi)} \right] \tilde{n}_q(t) + Re \left[ \sqrt{P} \sum_{k=-\infty}^{\infty} d_k \tilde{p}(t - kT) \frac{1}{j} e^{j(\Delta\omega t + \phi)} \right] \tilde{n}_i(t) \quad (A-8)$$

$$n'_{nn}(t) = \tilde{n}_i(t) \tilde{n}_q(t) \quad (A-9)$$

where the binary symbols  $d_k$  are  $\pm 1$  at random with probability one-half,  $\tilde{p}(t)$  and  $C_n(R)$  are given by Eqs. (B-8) and (B-9), and  $\tilde{n}_i(t)$  and  $\tilde{n}_q(t)$  are given by Eqs. (15) and (16). These noises are independent of one another since data and noise are assumed to be independent. Consequently, the autocorrelation of the effective noise is the sum of the autocorrelation of each of the noises. Mathematically,

$$R'_{eff}(\tau) = R'_{ss}(\tau) + R'_{sn}(\tau) + R'_{nn}(\tau) \quad (A-10)$$

where, from Appendix C,

$$R'_{ss}(\tau) \approx \frac{P^2 T}{8\pi^2 R^2} \left( -1 + \frac{e^{-2\pi R} + e^{2\pi R}}{2} \right)^2 \left( \frac{e^{-4\pi R}}{1 - e^{-4\pi R}} \right) \delta(\tau) \quad (A-11)$$

$$R'_{sn}(\tau) = \begin{cases} \frac{P}{2} \left[ 2 \left( 1 - \frac{|\tau|}{T} \right) - \frac{e^{-2\pi B|\tau|}}{2\pi R} (2 - e^{-2\pi R}) + \frac{e^{2\pi B|\tau|}}{2\pi R} e^{-2\pi R} \right] \frac{N_0}{2} \pi B e^{-2\pi B|\tau|}, & |\tau| \leq T \\ -\frac{P}{2} \frac{e^{-2\pi B|\tau|}}{2\pi R} (2 - e^{-2\pi R} - e^{2\pi R}) \frac{N_0}{2} \pi B e^{-2\pi B|\tau|}, & |\tau| \geq T \end{cases} \quad (A-12)$$

$$R'_{nn}(\tau) = \left( \frac{N_0}{2} \pi B \right)^2 e^{-4\pi B|\tau|} \quad (A-13)$$

The preceding equations describe the second-order statistics of the noise, which is usually fed back to close the loop along with the signal component. Consequently, the equations can be used to compute the tracking variance of a Costas loop with one-pole arm filters. The tracking variance was computed (in the absence of self-noise) by using the noise statistics above. It was found to be the same as that computed in Chapter 5 of [7]. Although this is not a definitive test of the accuracy of the equations above, it certainly enhances confidence in them.

As shown in Fig. 1, the error signal  $z'(t)$  is integrated and dumped every  $T$  sec to obtain the error sequence  $z(n)$ . That is,

$$z(n) = \frac{1}{T} \int_{nT}^{(n+1)T} z'(t) dt \quad (A-14)$$

Assuming that the frequency error to be estimated,  $\Delta f$ , is small relative to the data rate,  $R_s = 1/T$ , allows one to write  $z(n)$  as

$$z(n) = A(P, R) \sin[2\Delta\omega(nT + T/2) + 2\phi + \xi] + n_{eff}(n) \quad (\text{A-15})$$

where the sinusoids in Eq. (A-4), which are approximately constant over  $T$  sec by assumption, can be removed outside the integral after evaluating them at the midpoint of the integration window. The amplitude  $A(P, R)$  is given by Eq. (A-5), the phase  $\xi$  is not computed because it is not relevant, and the effective noise  $n_{eff}(n)$  is given by Eq. (21), where

$$n_{ss}(n) = \frac{1}{T} \int_{nT}^{(n+1)T} n'_{ss}(t) dt \quad (\text{A-16})$$

$$n_{sn}(n) = \frac{1}{T} \int_{nT}^{(n+1)T} n'_{sn}(t) dt \quad (\text{A-17})$$

$$n_{nn}(n) = \frac{1}{T} \int_{nT}^{(n+1)T} n'_{nn}(t) dt \quad (\text{A-18})$$

where  $n'_{ss}(t)$ ,  $n'_{sn}(t)$ , and  $n'_{nn}(t)$  are given by Eqs. (A-7) through (A-9). The discrete autocorrelation functions of the noises above, which are computed in Appendix C, are

$$R_{ss}(m) = \begin{cases} \frac{P^2}{8\pi^2 R^2} \left(-1 + \frac{e^{-2\pi R} + e^{2\pi R}}{2}\right)^2 \frac{e^{-4\pi R}}{1 - e^{-4\pi R}}, & m = 0 \\ 0, & m \neq 0 \end{cases} \quad (\text{A-19})$$

$$R_{sn}(m) = \begin{cases} P\sigma^2 \left[1 - \frac{1}{2\pi R} \left(\frac{5}{2} - \frac{9}{8\pi R}\right) + e^{-2\pi R} \left(\frac{1}{4} + \frac{1}{4(2\pi R)} - \frac{17}{8(2\pi R)^2}\right) - \frac{e^{-4\pi R}}{4(2\pi R)^2} + \frac{e^{-6\pi R}}{8(2\pi R)^2}\right], & m = 0 \\ P\sigma^2 \left[\frac{1}{4\pi R} \left(1 - \frac{9}{8\pi R}\right) + \frac{e^{-2\pi R}}{8} \left(1 + \frac{5}{2\pi R} + \frac{9}{4\pi^2 R^2}\right) - \frac{1}{8\pi R} \left(1 + \frac{1}{8\pi R}\right) (e^{-6\pi R} - 2e^{-8\pi R} + e^{-10\pi R})\right], & m = \pm 1 \\ P\sigma^2 \left(\frac{2 - e^{-2\pi R} + e^{2\pi R}}{64\pi^2 R^2}\right) (2e^{-4\pi R|m|} - e^{-4\pi R(|m|+1)} - e^{-4\pi R(|m|-1)}), & m \neq 0, \pm 1 \end{cases} \quad (\text{A-20})$$

$$R_{nn}(m) = \begin{cases} \sigma^4 \left(\frac{\pi R}{2} - \frac{1}{8}(1 - e^{-4\pi R})\right), & m = 0 \\ \frac{\sigma^4}{16} (e^{-4\pi R(|m|-1)} + e^{-4\pi R(|m|+1)} - 2e^{-4\pi R|m|}), & m \neq 0 \end{cases} \quad (\text{A-21})$$

where  $\sigma^2 = N_0/2T$  and  $R = BT$  is the ratio of the arm filter 3-dB bandwidth  $B$  to the data rate,  $1/T$ . Note that, as  $R \rightarrow \infty$ ,  $R_{ss}(m) \rightarrow 0$  as expected. In this case, the data pulses approach the unfiltered case and consequently, the self-noise power approaches zero. In the limit  $R \rightarrow \infty$ ,  $R_{sn}(m) \rightarrow P\sigma^2$ , which is the signal-noise power at the output of the integrate-and-dump filter when the input is white noise with spectral level  $PN_0/2$  watts<sup>2</sup>/Hz. Finally,  $R_{nn}(m) \rightarrow \infty$  as  $R \rightarrow \infty$ , since in this case the input is white noise with spectral level  $N_0^2\pi B/2$ , which becomes unbounded as  $R \rightarrow \infty$ ,  $R = BT$ , for fixed  $T$ .

## Appendix B

### Computing $\tilde{d}^2(t)$ in Eq. (A-1)

The data sequence  $d(t)$ , which is given mathematically by Eq. (2), has its Fourier transform given by

$$D(\omega) = \sum_{k=-\infty}^{\infty} d_k P(\omega) e^{-jk\omega T} \quad (\text{B-1})$$

where the binary symbols  $d_k$  are  $\pm 1$  with equal probability and  $P(\omega)$  is the Fourier transform, denoted  $\mathcal{F}\{\}$ , of the baseband NRZ pulse  $p(t)$ . Applying Eq. (B-1) to Eq. (10) one obtains

$$\tilde{D}(\omega) = \sum_{k=-\infty}^{\infty} d_k \tilde{P}(\omega) e^{-jk\omega T} \quad (\text{B-2})$$

where  $\tilde{P}(\omega)$  is defined to be

$$\tilde{P}(\omega) = P(\omega) H(\omega + \Delta\omega) \quad (\text{B-3})$$

As a result,  $\tilde{d}(t) = \mathcal{F}^{-1}\{\tilde{D}(\omega)\}$  is given as

$$\tilde{d}(t) = \sum_{k=-\infty}^{\infty} d_k \tilde{p}(t - kT) \quad (\text{B-4})$$

where  $\tilde{p}(t)$  is the inverse Fourier transform of  $\tilde{P}(\omega)$ . From Eq. (B-4), the term  $\tilde{d}^2(t)$  in Eq. (A-1) can be written as

$$\tilde{d}^2(t) = \sum_{k=-\infty}^{\infty} \tilde{p}^2(t - kT) + \sum_{k=-\infty}^{\infty} \sum_{\substack{l=-\infty \\ l \neq k}}^{\infty} d_k d_l \tilde{p}(t - lT) \tilde{p}(t - kT) \quad (\text{B-5})$$

The first term in Eq. (B-5) is a periodic complex function with period  $T$ . Consequently, it can be represented by a Fourier series:

$$\sum_{k=-\infty}^{\infty} \tilde{p}^2(t - kT) = \sum_{n=-\infty}^{\infty} C_n e^{j(2n\pi t/T)} \quad (\text{B-6})$$

where the Fourier coefficients  $C_n$  are given by

$$C_n = \frac{1}{T} \mathcal{F}\{\tilde{p}^2(t)\}_{\omega=n(2\pi/T)} \quad (\text{B-7})$$

For a one-pole filter,  $\tilde{p}(t)$  can be computed to be

$$\tilde{p}(t) = \frac{\omega_B}{\omega_B + j\Delta\omega} \left( p(t) - e^{-(\omega_B + j\Delta\omega)t} u(t) + e^{-(\omega_B + j\Delta\omega)(t-T)} u(t - T) \right) \quad (\text{B-8})$$

where  $\omega_B$  is the 3-dB filter bandwidth in radians,  $p(t) = u(t) - u(t - T)$ , and  $u(t)$  is the unit step function. Applying Eq. (B-8) to Eq. (B-7), the Fourier coefficients  $C_n$  as a function of  $\Delta\omega T$  and  $R = BT$  are computed to be

$$C_n(R) = \left( \frac{1}{1 + j \frac{\Delta\omega T}{2\pi R}} \right)^2 \left( (-1)^n \frac{\sin n\pi}{n\pi} - \frac{1}{2\pi R} \frac{(1 + j \frac{\Delta\omega T}{2\pi R}) (1 - e^{-(2\pi R + j\Delta\omega T)})}{(1 + j \frac{2n\pi + \Delta\omega T}{2\pi R}) (1 + j \frac{n\pi + \Delta\omega T}{2\pi R})} \right) \quad (\text{B-9})$$

Note that for the special cases of  $\Delta\omega = 0$  and  $\Delta\omega = (-n\pi)/T$ , Eq. (B-9) reduces to Table 2 of [9]. Additionally, for  $n = 0$ , Eq. (B-9) reduces to

$$C_0(R) = \left( \frac{1}{1 + j \frac{\Delta\omega T}{2\pi R}} \right)^2 \left( 1 - \frac{1}{2\pi R} \frac{1 - e^{-(2\pi R + j\Delta\omega T)}}{1 + j \frac{\Delta\omega T}{2\pi R}} \right) \quad (\text{B-10})$$

## Appendix C

### Computing $R'_{ss}(\tau)$ , $R'_{sn}(\tau)$ , $R'_{nn}(\tau)$ and $R_{SS}(m)$ , $R_{Sn}(m)$ , $R_{Nn}(m)$

When applicable, the derivations that follow assume that the frequency error  $\Delta f = 0$ . This greatly simplifies the computation of the various autocorrelation functions and is not expected to change appreciably the final result of the main text. The reason for this is that all the significant noise components give rise to continuous spectra whose shapes are not expected to change much with  $\Delta f$  when  $\Delta f \ll 1/T$ . Furthermore, when computing  $R'_{ss}(\tau)$ , the second term in Eq. (A-7) is not included. This approximation is made because, except for the first harmonic which has a line component at  $1/T \pm 2\Delta f$ , the remaining harmonics give rise to line components far beyond the band of frequencies within which the error signal is expected. Additionally, power in the first harmonic which is at least 8 dB down from the component at  $2\Delta f$  when  $\Delta\omega T \leq 0.1$ , is suppressed even further by the integrate-and-dump filter response preceding the FFT operation.

#### I. $R'_{ss}(\tau)$ and $R_{SS}(m)$

The self-noise  $n'_{s,s}(t)$  is given by Eq. (A-7). When  $\Delta\omega = 0$ ,  $\tilde{p}(t)$  is real, and  $n'_{s,s}(t)$  reduces to

$$n'_{s,s}(t) = \frac{P}{2} \left( \sum_{k=-\infty}^{\infty} \sum_{\substack{l=-\infty \\ l \neq k}}^{\infty} d_k d_l \tilde{p}(t - lT) \tilde{p}(t - kT) \right) \sin 2\phi \quad (\text{C-1})$$

where  $d_n$  is  $\pm 1$  at random with probability one-half, and  $\phi$ , which is independent of  $d_n$ , is a uniform random variable in the interval  $[0, 2\pi]$ . The filtered pulse  $\tilde{p}(t)$  is given by Eq. (B-8). The autocorrelation function of  $n'_{s,s}(t)$  is given by

$$R'_{s,s}(\tau) = \mathcal{E}[n'_{s,s}(t)n'_{s,s}(t + \tau)] = \frac{P^2}{8} R(\tau) \quad (\text{C-2})$$

where, after noting that  $\mathcal{E}[\sin^2 \phi] = \frac{1}{2}$ ,  $R(\tau)$  can be defined as follows:

$$R(\tau) = \sum_{k=-\infty}^{\infty} \sum_{\substack{l=-\infty \\ l \neq k}}^{\infty} \sum_{\substack{m=-\infty \\ m \neq n}}^{\infty} \sum_{\substack{n=-\infty \\ n \neq m}}^{\infty} d_k d_l d_m d_n \tilde{p}(t - kT) \tilde{p}(t - lT) \tilde{p}(t - mT) \tilde{p}(t - nT) \quad (\text{C-3})$$

The function above, which is in general difficult to compute because the filtered pulse  $\tilde{p}(t)$  is not time limited, can be approximated as follows. One begins by expressing  $R(\tau)$  as the inverse Fourier transform of its corresponding spectral density.

$$R(\tau) = \int_{-\infty}^{\infty} S(\omega) e^{j\omega\tau} \frac{d\omega}{2\pi} \approx S(0) \lim_{\omega \rightarrow 0} \int_{-\omega}^{\omega} e^{j\omega\tau} \frac{d\omega}{2\pi} \quad (\text{C-4})$$

From Chapter 5 of [6], the dc component of the spectral density in Eq. (C-4) is given by

$$S(0) = 4T \sum_{l=1}^{\infty} R_d^2(lT) \quad (\text{C-5})$$

where

$$R_d(lT) = \int_{-\infty}^{\infty} S_d(\omega) |H(\omega)|^2 e^{-j\omega lT} \frac{d\omega}{2\pi} \quad (\text{C-6})$$

where  $S_d(\omega)$  is the spectral density of the random NRZ data and  $|H(\omega)|^2$  is the squared magnitude of the one-pole filter with 3-dB bandwidth  $B$  Hz. Grouping the exponential in the above integral with  $|H(\omega)|^2$  and applying Parseval's theorem gives the corresponding integral in the transformed domain. Namely,

$$R_d(lT) = \int_{-T}^T \left(1 - \frac{|\tau|}{T}\right) \pi B e^{-2\pi B|\tau+lT|} d\tau = \frac{e^{-2\pi Rl}}{2\pi R} \left(-1 + \frac{e^{-2\pi R} + e^{2\pi R}}{2}\right) \quad (\text{C-7})$$

where the parameter  $R = BT$ . Finally,  $R'_{ss}(\tau)$  in Eq. (A-11) follows after back-substituting the last result.

The discrete autocorrelation function given by Eq. (A-19), which is the correlation function of the noise sequence  $n_{ss}(n)$  given by Eq. (A-16), follows directly upon evaluating the next equation:

$$R_{ss}(m) = \frac{1}{T^2} \int_0^T \int_{mT}^{(m+1)T} R'_{ss}(|t_2 - t_1|) dt_2 dt_1 \quad (\text{C-8})$$

where  $R'_{ss}(\tau)$  is given by Eq. (A-11).

## II. $R'_{sn}(\tau)$ and $R_{sn}(m)$

The signal-noise process  $n'_{sn}(t)$  is given by Eq. (A-8). For  $\Delta\omega = 0$ ,  $\tilde{p}(t)$  is real and  $n'_{sn}(t)$  reduces to

$$n'_{sn}(t) = \left(\sqrt{P} \sum_{k=-\infty}^{\infty} d_k \tilde{p}(t - kT) \cos \phi\right) \tilde{n}_q(t) + \left(\sqrt{P} \sum_{k=-\infty}^{\infty} d_k \tilde{p}(t - kT) \sin \phi\right) \tilde{n}_i(t) \quad (\text{C-9})$$

Since the data and noise processes are independent of each other, the autocorrelation of the signal-noise product is the product of the autocorrelation of the individual signal and noise processes. Hence,

$$R'_{sn}(\tau) = \mathcal{E}[n'_{sn}(t)n'_{sn}(t + \tau)] = PR_s(\tau)R_n(\tau) \quad (\text{C-10})$$

where

$$R_s(\tau) = \mathcal{E} \left[ \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} d_k d_l \tilde{p}(t - kT) \tilde{p}(t - lT) \right] \quad (\text{C-11})$$

$$R_n(\tau) = \mathcal{E}[\tilde{n}_i(t)\tilde{n}_i(t + \tau)] = \mathcal{E}[\tilde{n}_q(t)\tilde{n}_q(t + \tau)] \quad (\text{C-12})$$

The functions  $R_s(\tau)$  and  $R_n(\tau)$  are the autocorrelation functions of signal and noise processes at the output of a one-pole filter when the input signal process is random NRZ data, and the input noise process is white noise. Consequently, writing  $R_s(\tau)$  as the inverse Fourier transform of its corresponding spectral density in terms of the input spectral density times the squared magnitude of the one-pole filter, grouping the exponential with the squared magnitude of the filter transfer function, and applying Parseval's theorem yields

$$R_s(\tau) = \int_{-T}^T \left(1 - \frac{|\tau'|}{T}\right) \frac{\pi B}{2} e^{-2\pi B|\tau'+\tau|} d\tau' \quad (\text{C-13})$$

where  $B$  Hz is the 3-dB bandwidth of the one-pole filter. The noise correlation is well known to be

$$R_n(\tau) = \frac{N_0}{2} \pi B e^{-2\pi B|\tau|} \quad (\text{C-14})$$

Finally, the signal–noise correlation given by Eq. (A-12) follows after substituting Eqs. (C-13) and (C-14) into Eq. (C-10).

The discrete autocorrelation function given by Eq. (A-20), which is the correlation function of the noise sequence  $n_{s,n}(n)$ , is obtained upon evaluating the next equation:

$$R_{s,n}(m) = \frac{1}{T^2} \int_0^T \int_{mT}^{(m+1)T} R'_{s,n}(|t_2 - t_1|) dt_2 dt_1 \quad (\text{C-15})$$

Since the integrand above depends only on the absolute difference  $|t_2 - t_1|$ , the double integral above can be reduced to the following simpler single integrals when  $m = 0, 1, -1$ . Namely,

$$\int_0^T \int_{mT}^{(m+1)T} R'_{s,n}(|t_2 - t_1|) dt_2 dt_1 = \begin{cases} \int_{-T}^T R'_{s,n}(|\tau|)(T - |\tau|) d\tau, & m = 0 \\ \int_0^{2T} R'_{s,n}(|\tau|) \tau d\tau, & m = 1, -1 \end{cases} \quad (\text{C-16})$$

When  $|m| > 1$ , Eq. (C-15) reduces to the product of two single integrals in a straightforward manner.

### III. $R'_{nn}(\tau)$ and $R_{nn}(m)$

The noise process  $n'_{nn}(t)$  given by Eq. (A-9) is the product of the filtered inphase and quadrature noise processes,  $\tilde{n}_i(t)$  and  $\tilde{n}_q(t)$ . The autocorrelation of these noises, denoted by  $R_n(\tau)$ , is given by Eq. (C-14). Furthermore, since  $\tilde{n}_i(t)$  and  $\tilde{n}_q(t)$  are independent with respect to each other,  $R_{nn}(\tau) = R_n^2(\tau)$  and Eq. (A-13) follows. The discrete autocorrelation  $R_{nn}(m)$  is given upon evaluating the following equation:

$$R_{nn}(m) = \int_0^T \int_{mT}^{(m+1)T} R'_{nn}(|t_2 - t_1|) dt_2 dt_1 \quad (\text{C-17})$$

where, when  $m = 0$ , the double integral above can be transformed to the following simpler single integral:

$$\int_0^T \int_0^T R'_{nn}(|t_2 - t_1|) dt_2 dt_1 = \int_{-T}^T R'_{nn}(|\tau|) (T - |\tau|) d\tau \quad (\text{C-18})$$

When  $m \neq 0$ , Eq. (C-17) reduces to the product of two single integrals.

## References

- [1] S. Hinedi, "A Functional Description of the Advanced Receiver," *TDA Progress Report 42-100*, vol. October–December 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 131–149, February 15, 1990.
- [2] B. Shah and S. Hinedi, "A Comparison of Open-Loop Suppressed-Carrier Acquisition Techniques," *TDA Progress Report 42-103*, vol. July–September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 170–188, November 15, 1990.
- [3] S. Aguirre, "Acquisition Times of Carrier Tracking Sampled Data Phase-Locked Loops," *TDA Progress Report 42-84*, vol. October–December 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 88–93, February 15, 1986.
- [4] S. Hinedi and B. Shah, "QPSK Carrier-Acquisition Performance in the Advanced Receiver II," *TDA Progress Report 42-100*, vol. October–December 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 150–159, February 15, 1990.
- [5] D. C. Rife and R. R. Boorstyn, "Single-Tone Parameter Estimation From Discrete-Time Observations," *IEEE Trans. Info. Theory*, vol. IT-20, no. 5, pp. 591–598, September 1974.
- [6] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, New Jersey: Prentice-Hall, 1989.
- [7] J. K. Holmes, *Coherent Spread Spectrum Systems*, New York: John Wiley & Sons, 1982.
- [8] F. M. Gardner, *Phase Lock Techniques*, 2nd edition, New York: John Wiley & Sons, 1979.
- [9] G. Hedin, J. K. Holmes, W. C. Lindsey, and K. T. Woo, "False Lock Phenomenon in Costas and Squaring Loops," *Proc. 1977 National Telecommunications Conference*, pp. 30:1-1–30:1-5, Los Angeles, California, 1977.

97-32  
3-1-15  
N91-18325

JJ57 (107)

# Frame Synchronization for the Galileo Code

S. Arnold and L. Swanson  
Communications Systems Research Section

*This article reports results on the performance of the Deep Space Network's frame synchronizer for the (15,1/4) convolutional code after Viterbi decoding.*

## I. Introduction

Most deep-space mission communication systems use a 32-bit unique word (synchronization or sync word) to identify the beginning of a telemetry frame. The sync word is incorporated into the spacecraft's telemetry data stream after it has been divided into frames. The data stream is then encoded, transmitted, and received at a Deep Space Network (DSN) tracking station. The ability of the tracking station to achieve frame synchronization requires the appearance of the sync word within the decoded bit stream. Difficulty arises when the sync word has been corrupted by bit errors, which the Viterbi decoder tends to create in bursts. As a result, the DSN has incorporated frame synchronizers into its communication systems based on an algorithm that attempts to find the sync word even if it is corrupted.

The performance of the DSN frame synchronization algorithm has been analyzed for various frame lengths at various signal-to-noise ratios (SNRs) and thresholds using the NASA (7,1/2) convolutional code [1]. The (15,1/4) experimental convolutional code developed for the Galileo mission to Jupiter will use the same frame synchronizers. Although the present DSN frame synchronization scheme is adequate for the (7,1/2) code, its performance for the (15,1/4) code is unknown. Given the size of the Galileo

code, the average error bursts generated from the output of the (15,1/4) Viterbi decoder are twice as long as the bursts from the (7,1/2) Viterbi decoder. See [2,3] for further discussion of the burst statistics for the NASA code and the Galileo code. This article determines the performance of the frame synchronizer for the (15,1/4) convolutional code after Viterbi decoding and finds the threshold that optimizes the probability of acquiring true sync within four frames using a strategy that requires next-frame verification.

## II. Method for Finding Sync

Many different frame synchronization methods have been studied, but the DSN's current method is to compare the true sync word (noiseless version) to a 32-bit segment of decoded bits. Those bits found to be in disagreement are counted. This count is then compared to a predetermined threshold,  $T$ , optimized for a given bit error rate (BER). If the number of disagreements is greater than  $T$ , those 32 bits are rejected as the sync word. Otherwise, the 32 bits are recorded as a sync word candidate. Successive one-bit shifts of 32-bit received signal segments (sliding window) are compared to the true sync word until the threshold test is passed at the same location in two consecutive frames. Once sync has been declared, testing for the sync word

continues through all succeeding frames. If sync is lost, the sliding window process is started again.

### III. Analysis

Decoded data bits were generated using the 1-kbps Viterbi decoder [4], also known as the Little Viterbi Decoder (LVD). The LVD is a hardware decoder developed for use in testing long constraint-length convolutional codes. The received symbols fed into the decoder represent encoded symbols generated from the all-zero information bit sequence with pseudo-random noise added (i.e., noise only). The LVD generated enough data to ensure that 100 error bursts were produced for each SNR of interest. For each SNR tested, the decoded bits were subjected to the threshold test for possible threshold values  $T$ , where  $0 \leq T \leq 10$ . From a random 32-bit window of decoded bits, this test determines whether the number of decoded bit errors in the observed window exceeds the given threshold. A count is maintained of the number of 32-bit windows for which the number of errors exceeds  $T$ . The 32-bit window is then shifted to the right one bit until all possible 32-bit segments have been tested.

The two error components that influence the overall performance of the frame synchronization scheme are the probability of miss  $P_m$  and the probability of false alarm  $P_f$ .  $P_m$  is the likelihood that the sync word is not detected in the decoded bit stream.  $P_f$  is the likelihood that the sync word is falsely detected in an incorrect position in the decoded bit stream. For this article,  $P_m$  was estimated from the LVD error data as

$$P_m = \frac{X}{Y} \quad (1)$$

where

$X$  = number of 32-bit windows where the number of errors exceeds  $T$

$Y$  = number of 32-bit windows tested within a given file

Assuming random data,  $P_f$  is given by [5]

$$P_f = \sum_{k=0}^T \binom{32}{k} 2^{-32} \quad (2)$$

Note that  $P_m$  depends on the code and the SNR, but  $P_f$  does not.

Failure to acquire sync in one frame will occur if either the location of the true sync word fails the threshold test or if the sync word is falsely detected before the location of the true sync word. The probability of acquiring sync correctly within one frame  $P_C$  can be approximated for small values of  $P_m$  and  $P_f$  by [6]

$$P_C \approx 1 - P_m - \frac{B-1}{2} P_f \quad (3)$$

where

$$B = \text{length of data frame}$$

Next-frame verification requires that the 32-bit sync word candidate found in the current frame be verified in the next succeeding frame. To acquire sync correctly within four frames, the threshold test must be passed correctly in one of three ways: within the first and second frames; in the second and third frames after failing in the first frame; or in the third and fourth frames after failing in the second frame. Equation (3) can be extended to the probability of acquiring sync correctly within four frames with next-frame verification  $P_{C4}$  for small values of  $P_m$  and  $P_f$  by [6]

$$P_{C4} \approx 1 - 3P_m^2 - \frac{B-1}{2} P_f^2 \quad (4)$$

assuming that the decoded bit errors in the four frames tested are independent.

### IV. Results

If the SNR over the DSN channel were sufficiently strong to ensure no bit errors, the received (decoded) sync word would be identical to the true sync word. However, no matter what the SNR and threshold are, there is always a nonzero probability that the sync word will not be found correctly due to the possibility of random data mimicking the sync word within  $T$  or fewer disagreements. Given this possibility, the choice of threshold  $T$  requires a trade-off between  $P_m$  and  $P_f$ . As  $T$  increases,  $P_m$  improves very little while  $P_f$  increases substantially. This effect can be seen in Fig. 4. Figure 5 shows how individual error components  $P_m$  and  $P_f$  work to influence the overall performance of the frame synchronization scheme.

Several figures have been drawn to quantitatively describe sync-acquisition probability using the (15,1/4) convolutional code. In all cases,  $P_m$  is determined from the output data of the LVD;  $P_f$  is calculated from Eq. (2); and

the overall probability of acquiring sync correctly within four frames is calculated from Eq. (4). The sync marker length is 32 bits. Figures 1 and 2 plot the probability of not finding sync correctly within four frames with next-frame verification for a frame length of 5120 bits with increasing SNR and  $T$ , respectively. Figures 3 and 4 plot the same data for a frame length of 8960 bits. Figure 4 also shows a limiting case,  $\text{SNR} = \infty$  ( $P_m = 0$ ), for which the probability of not correctly acquiring sync is entirely due to the possibility of false alarms. If threshold  $T$  is set too high, the result is unacceptable performance no matter how high the SNR. Figure 5 replots one of the curves from Fig. 4 with additional curves showing two individual components ( $3P_m^2$  and  $\frac{B-1}{2}P_f^2$ ) contributing to the probability of not correctly acquiring sync. In this figure,  $P_f$  is shown to have little or no influence on the performance until the critical point where  $P_m$  and  $P_f$  intersect ( $T \approx 6$ ). At this point  $P_f$  begins to overwhelm  $P_m$ . Note that the  $P_m$  component is dominant for smaller thresholds ( $T \leq 5$ ) while the  $P_f$  component is dominant for higher thresholds ( $T \geq 6$ ). Figure 6 plots  $P_m$  versus  $P_f$  curves for various SNRs. The plot symbols in this figure represent threshold values  $T$  when  $0 \leq T \leq 7$ . Observe that the

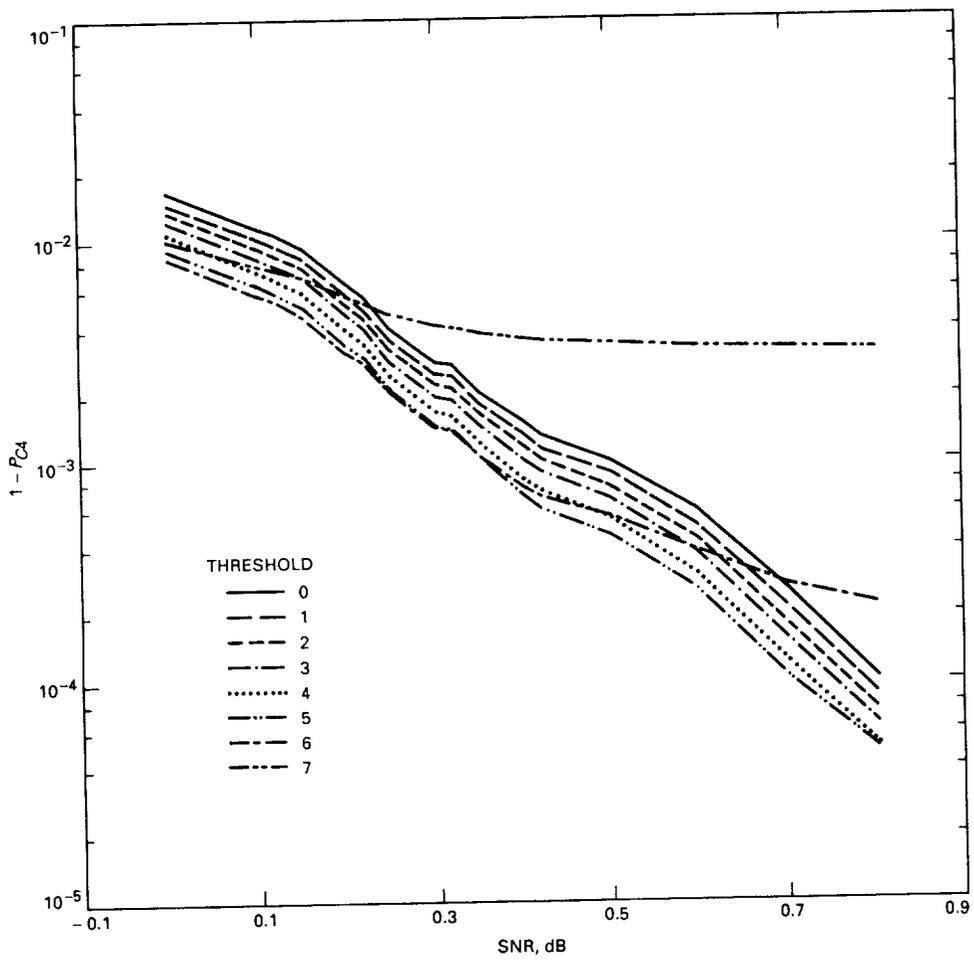
corresponding thresholds of each curve have the same  $P_f$  value, providing evidence as shown in Eq. (2) that  $P_f$  is independent of SNR.

## V. Conclusion

In general, choosing a threshold for the DSN communication system should be based on the system's operating point. However, other important factors to be considered include the frame-to-frame verification strategy, the length of a frame, and the size of the sync word. Therefore, the following recommendations are based only on the frame synchronization scheme described above. In order to maximize the probability of acquiring sync correctly within four frames with next-frame verification in the area of interest,  $\text{SNR} = 0.5$  dB (Viterbi decoder BER =  $5 \times 10^{-3}$ ), a threshold value of five would be optimal for each of the frame lengths tested (5120 and 8960 bits). However, using a threshold of five for SNRs above those tested would result in a sync acquisition rate that is less than optimal. If slightly higher SNRs are anticipated, a threshold of four would be more appropriate.

## References

- [1] C. S. R. Aguilera, L. Swanson, and G. H. Pitt, "Frame Synchronization Performance and Analysis," *TDA Progress Report 42-92*, vol. October-December 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 110-116, February 15, 1988.
- [2] R. L. Miller, L. J. Deutsch, and S. A. Butman, "On the Error Statistics of Viterbi Decoding and the Performance of Concatenated Codes," JPL Publication 81-9, Jet Propulsion Laboratory, Pasadena, California, 1981.
- [3] K.-M. Cheung and S. J. Dolinar, "Performance of Galileo's Concatenated Codes with Nonideal Interleaving," *TDA Progress Report 42-95*, vol. July-September 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 148-152, November 15, 1988.
- [4] C. R. Lahmeyer and K.-M. Cheung, "Long Decoding Runs for Galileo's Convolutional Codes," *TDA Progress Report 42-95*, vol. July-September 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 143-146, November 15, 1988.
- [5] L. Swanson, *A Comparison of Frame Synchronization Methods*, JPL Publication 82-100, Jet Propulsion Laboratory, Pasadena, California, 1982.
- [6] S. J. Dolinar and K.-M. Cheung, "Frame Synchronization Methods Based on Channel Symbol Measurements," *TDA Progress Report 42-98*, vol. April-June 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 121-137, August 15, 1989.



**Fig. 1. The effect of an increasing SNR on the probability of not correctly acquiring sync within four frames with next-frame verification, assuming a frame length of 5120 bits.**

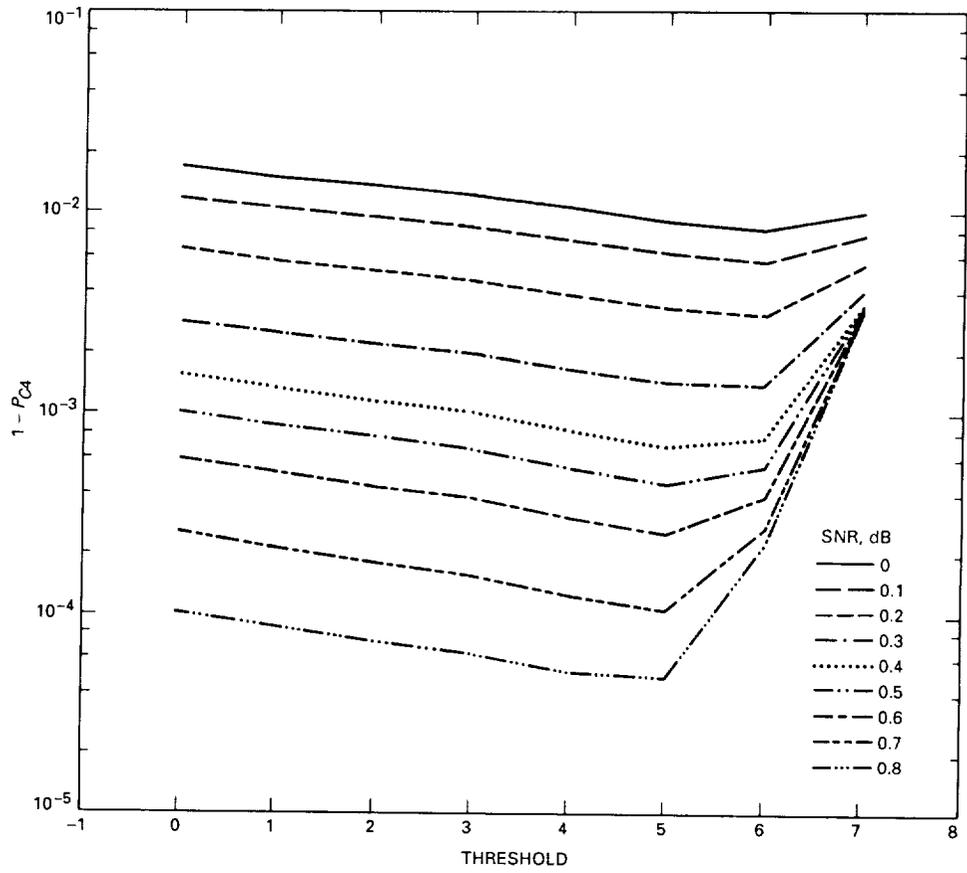


Fig. 2. The effect of threshold on the same probability as described in Fig. 1.

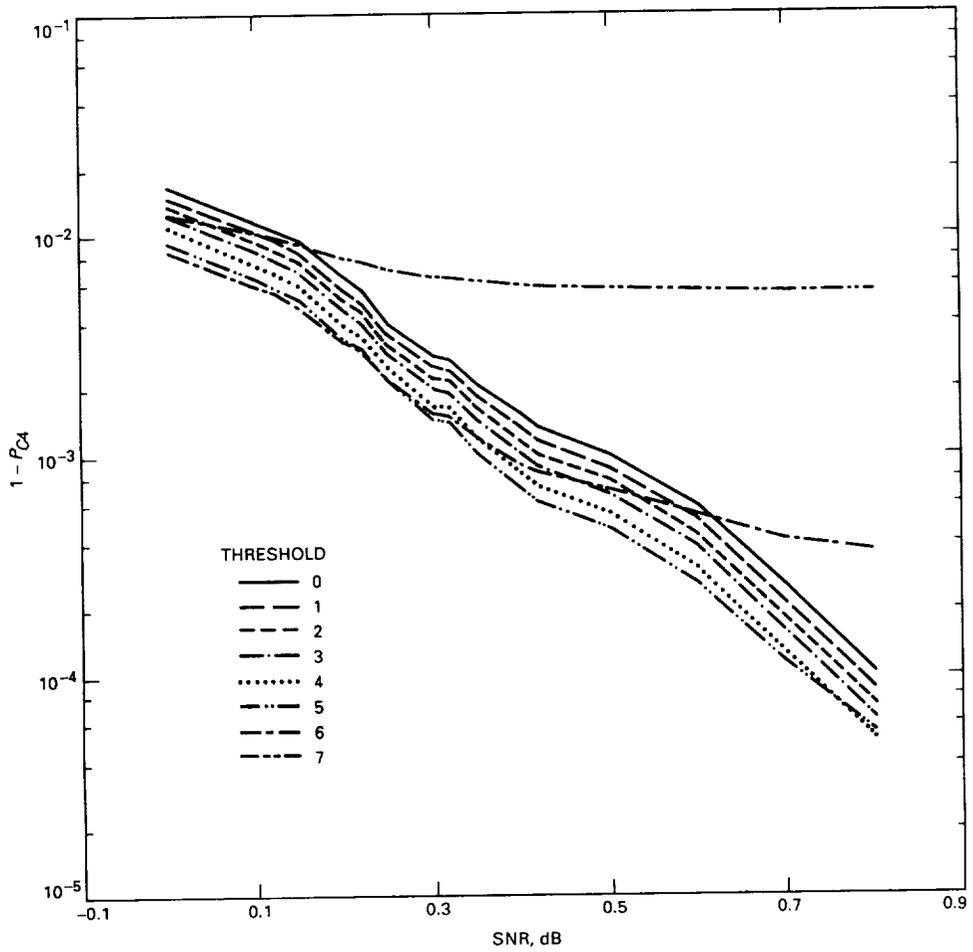


Fig. 3. The effect of SNR on the same probability as Figs. 1 and 2, assuming a frame length of 8960 bits.

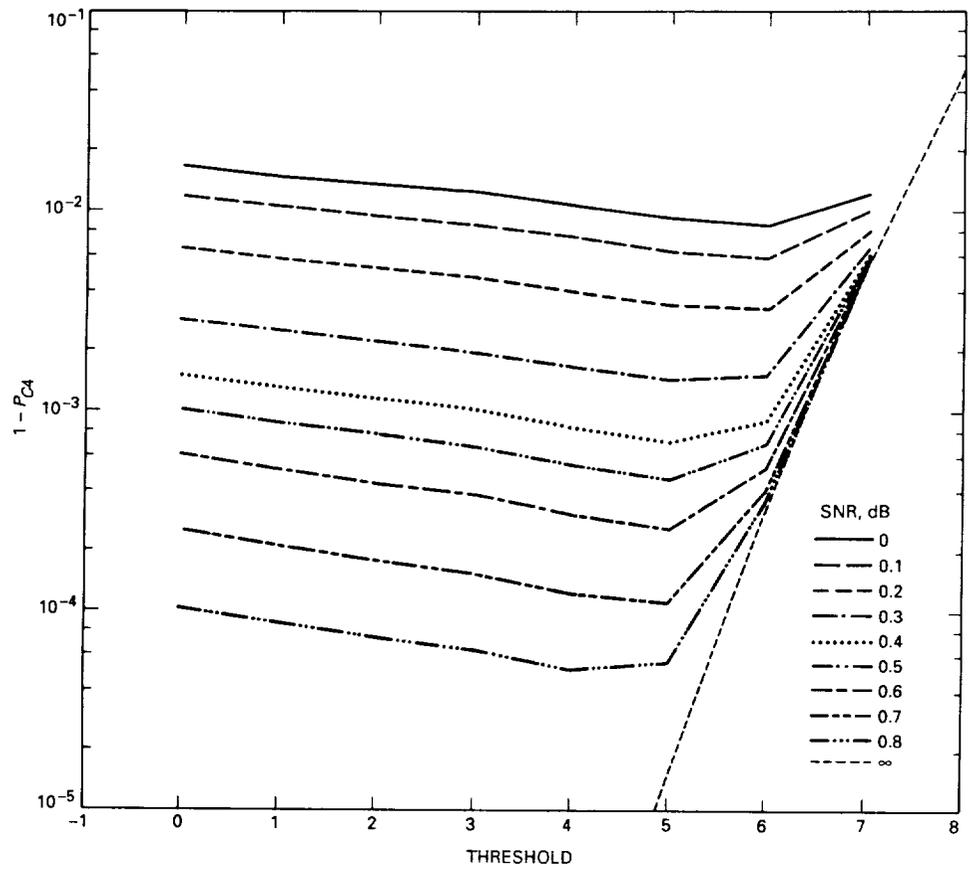


Fig. 4. The effect of threshold on the same probability as Fig. 3.

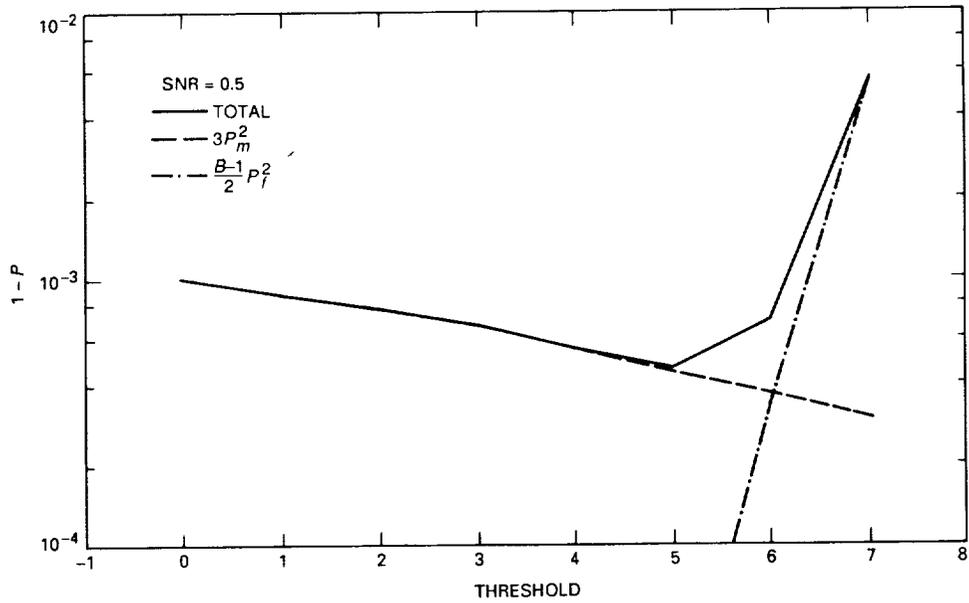


Fig. 5. The effect of two additional error factors to the probability described in Fig. 4.

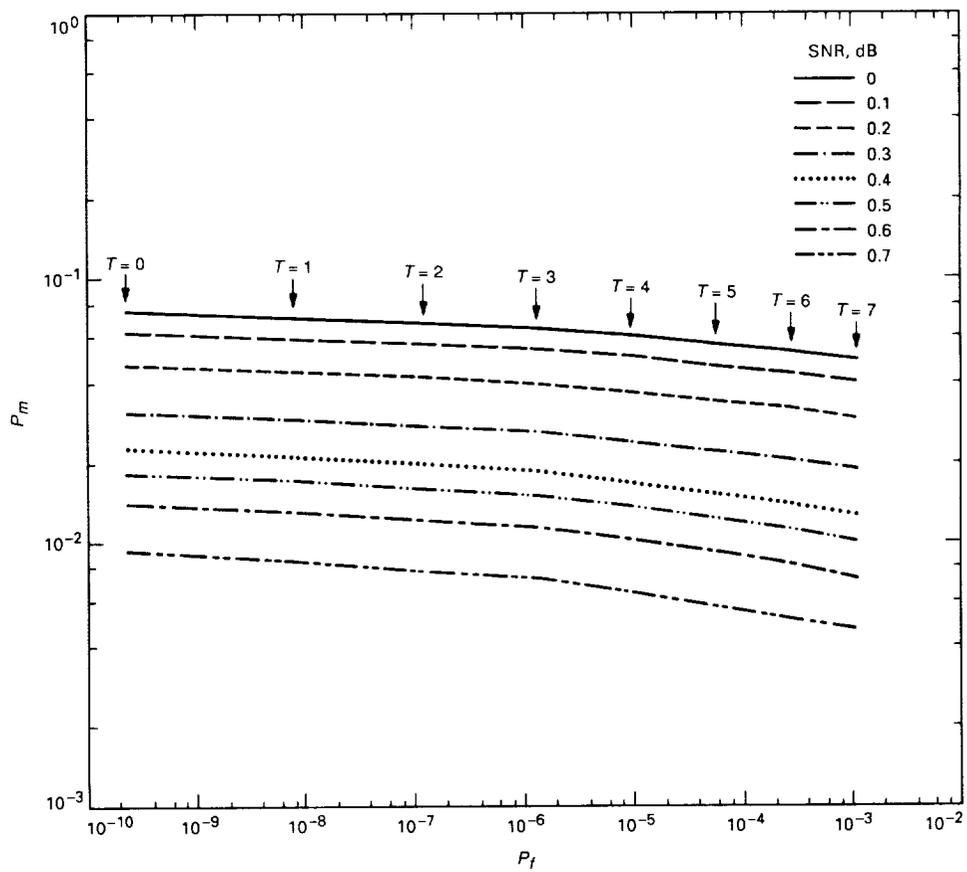


Fig. 6. The probability of miss  $P_m$  versus the probability of false alarm  $P_f$  for various SNRs.

518-89  
372076  
P-16  
N91-18326  
JJ57440

# Planetary Ephemerides Approximation for Radar Astronomy

R. Sadr and M. Shahshahani  
Communications Systems Research Section

*In this article the planetary ephemerides approximation for radar astronomy is discussed, and, in particular, the effect of this approximation on the performance of the programmable local oscillator (PLO) used in Goldstone Solar System Radar is presented. Four different approaches are considered and it is shown that the Gram polynomials outperform the commonly used technique based on Chebyshev polynomials. These methods are used to analyze the mean square, the phase error, and the frequency tracking error in the presence of the worst-case Doppler shift that one may encounter within the solar system. It is shown that in the worst case the phase error is under one degree and the frequency tracking error less than one hertz when the frequency to the PLO is updated every millisecond.*

## I. Introduction

Planetary ephemerides are used in radar astronomy to transmit a coherent beam in the direction of a planet. This beam is reflected from the surface of a planet, and the measured Doppler shift from the reflected beam is used to reconstruct a two-dimensional radar image [1,2]. Both the transmitter and the receiver frequencies may be programmed in some situations. For example, it is common to transmit the uplink signal to cause the frequency at one station to remain constant while correcting the frequencies at the other stations to compensate for the Doppler difference.

The Navigation Systems Section of the Jet Propulsion Laboratory provides high precision planetary and celestial body ephemerides for various studies. The ephemerides

are computed by numerical integration of a model of the solar system.<sup>1</sup> Saving the ephemerides at every integration step would result in prohibitively large data files, and it is not computationally feasible to run this program in real time to generate the ephemerides data. Thus, it is essential to approximate the ephemerides with a set of polynomials and use this set to generate the ephemerides in real time.

The main purpose of this article is to investigate four different approaches for the ephemerides approximation. It is concluded that the Gram polynomials consistently outperform the commonly used technique based on the Chebyshev polynomials. In fact, in some cases the mean-

<sup>1</sup> E. M. Standish, Jr. and D. K. Yeomans, Navigation Systems Section, Jet Propulsion Laboratory, Pasadena, California, private communication, November 14, 1989.

square error (MSE) is lower by a factor of one hundred. Furthermore, it is shown here that using a piecewise orthonormal expansion is superior to the classical least square fit when the number of data points is large. The computational complexity of using the Gram polynomials is equivalent to that of any other polynomial approximation of the same degree. One can intuitively explain the superior performance of the Gram polynomials by noting that they form a complete orthogonal set on an evenly partitioned interval (see Section IV.A).

For applications to NASA's Deep Space Network (DSN) and the Goldstone Solar System Radar (GSSR) [1,3], the ephemerides are converted into an integer valued frequency control word, which is used by a digital frequency synthesizer (DFS) to produce a sinusoid at an intermediate frequency (IF).<sup>2</sup> The theory and the design of the DFS for radar astronomy are described in detail in [4]. In Section VII, it is assumed that the reader is familiar with the theory of operation of the DFS as described in [4].

Figure 1 shows the overall configuration of the X-band exciter [5]. This exciter is used for the transmission of a coherent X-band signal from the DSN station to a distant planet. The reflected signals are used for generating radar images of the planet [1,2].

The design of the exciter is based on using a high-resolution programmable local oscillator (PLO) with controllable phase and frequency. The output of the PLO is ideally a single carrier with a frequency range of 10 to 20 MHz.<sup>2</sup> The block diagram of the PLO is given in Fig. 2. It is composed of a DFS and digital-to-analog conversion module, and is controlled by a host via a parallel interface.

The host software driver for the PLO controls both the frequency and the phase of the DFS. The phase and the frequency of the DFS are computed from the ephemerides data and are updated at a constant rate by the host. In approximating the ephemerides, it is important to use the least degree polynomial which gives the satisfactory approximation. As the degree of the approximating polynomial increases, the computation time also increases, and as a result fewer updates from the host will be possible. Furthermore, the effect of changing the frequency of the DFS as a function of the ephemerides introduces frequency modulation at the output of the PLO. In Section VII of this article both of these issues are addressed: namely, the effects of the update rate on the phase and frequency tracking error, and on the output spectrum of the PLO.

<sup>2</sup> F. R. Jurgens, "High Level PLO Definition," JPL Interoffice Memorandum 331-90.10-009 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 19, 1990.

## II. Effect of the Doppler Shift on the DFS Input

In the application of DFS for the GSSR, frequency control word  $F_r$  is updated at the fixed rate corresponding to the worst case Doppler shift that one may encounter. Let  $R$  denote the distance between the Earth and the planet that is being tracked. The rate of the change of this distance, when the planet is at position  $\mathbf{x} = \mathbf{x}(t)$ , is

$$\frac{dR}{dt} = \frac{\mathbf{v} \cdot \mathbf{x}}{R} \quad (1)$$

where  $\mathbf{v}$  denotes the velocity of the planet relative to the Earth. Let  $F_o$  denote the output frequency, and  $c$  the speed of light, then the Doppler shift  $\Delta F$  is

$$\Delta F = \frac{dR/dt}{cF_o} \quad (2)$$

The ephemerides provide the value of  $\Delta F$  in tabular form, and as explained earlier, it is essential to fit a polynomial to the function  $\phi(t) = dR/dt$ .

## III. Method of Least Squares

In approximating the ephemerides, it is assumed that there exists a real valued function  $f(x)$  such that  $f(x_i)$  represents the sampled values  $f_i$  at the point  $x_i$ . The problem of approximating the sequence  $\{f_i, i = 1, \dots, N\}$ , with a function  $y(x)$ , is formulated by using the MSE measure as a merit function [2,6]. A space of functions is fixed and a basis  $\Phi = \{\phi_i\}$  is supplied. Let  $\chi$  denote the MSE, i.e.,

$$\chi^2 = \sum_{i=1}^N \left( f(x_i) - \sum_{k=1}^M a_k \phi_k(x_i) \right)^2 \quad (3)$$

The coefficients  $a_1, \dots, a_M$  are chosen to minimize  $\chi^2$ . For such  $a_i$ , set

$$y(x) = \sum_{k=1}^M a_k \phi_k(x) \quad (4)$$

For example, if  $\phi_i(x) = x^{i-1}$  with  $i = 1, \dots, M$ , then a polynomial approximation to  $f$  is obtained.

If one defines the  $N \times M$  matrix  $A$  and the  $N \times M$  vector  $\mathbf{b}$  by

$$A = \begin{pmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_M(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_M(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_N) & \phi_2(x_N) & \cdots & \phi_M(x_N) \end{pmatrix},$$

$$\mathbf{b} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{pmatrix} \quad (5)$$

then the solution  $\mathbf{a} = (a_1, \dots, a_M)$  to the above minimization problem satisfies the equation

$$A^T A \mathbf{a} = A^T \mathbf{b} \quad (6)$$

This equation is also known as the normal equation of the least squares. It is convenient to set  $\Lambda = (\lambda_{kj})$  and  $B = (\beta_k)$  where

$$\lambda_{kj} = \sum_{i=1}^N \phi_j(x_i) \phi_k(x_i)$$

$$\beta_k = \sum_{i=1}^N f_i \phi_k(x_i) \quad (7)$$

Then Eq. (6) becomes

$$\Lambda \mathbf{a} = B \quad (8)$$

In most applications (especially when  $N$  is large), the normal equation is nearly singular, and singular value decomposition (SVD) must be employed [7] to solve for  $\mathbf{a}$ . SVD requires much more extra storage and computation than solving the normal equation. In Section IV, an alternative solution for this minimization problem is considered.

#### IV. Orthonormal Expansion

In this section it is assumed that  $\{\phi_i(x), i = 1, 2, \dots\}$  is a complete orthonormal set relative to a measure  $d\mu(x)$ , in the interval  $[a, b]$ , i.e.,  $\langle \phi_i(x), \phi_j(x) \rangle = \delta_{ij}$ , for all  $i \neq j$ , where the inner product  $\langle \cdot \rangle$  is given by

$$\langle y(x), z(x) \rangle = \int_a^b y(x)z(x)d\mu \quad (9)$$

Then  $y(x) = \sum \langle y, \phi_j \rangle \phi_j(x)$  in the  $L^2$ -sense.

Any function  $f(x) \in C_{[a,b]}$ , where  $C_{[a,b]}$  denotes the space of continuous functions on the interval  $[a, b]$ , has an approximate expansion

$$f(x) \approx \sum_{i=1}^M a_i \phi_i(x) \quad (10)$$

where  $M$  is large. An exact solution is, in general, not possible, and the choice of  $\mathbf{a} = (a_1, \dots, a_M)$  which minimizes

$$\chi^2 = \int_a^b \left| f(x) - \sum_{i=1}^M a_i \phi_i(x) \right|^2 d\mu \quad (11)$$

is given by  $a_i = \langle f, \phi_i \rangle$ . Similar considerations apply to the discrete case. Here the domain of the functions is the set  $\Delta = \{x_1, x_2, \dots, x_N\}$ , and  $\mu$  is a non-negative measure on  $\Delta$ . It is assumed that the points  $x_i$  are equally spaced in the interval  $[a, b]$ . The inner product becomes

$$\langle y(x), z(x) \rangle = \sum_{i=1}^N y(x_i)z(x_i)\mu(x_i) \quad (12)$$

When the basis functions form a complete orthonormal set, then all the nondiagonal terms in the matrix  $\Lambda$  are zero and the computation of  $\mathbf{a}$  in Eq. (8) is reduced to inverting the diagonal matrix  $\Lambda$ . Hence, the complexity is substantially reduced when compared to directly solving the simultaneous set of normal equations in Eq. (8), or using the SVD method. This savings is accomplished without any degradation in the overall average MSE. In the following two sections, two special classes of orthogonal polynomial equations that are used for computing the interpolating polynomial for the ephemerides are specifically considered.

#### A. Chebyshev Polynomial Equations

The Chebyshev polynomials have been widely used for approximating the planetary ephemerides [3]. In this section, this class of polynomials is described and its shortcoming for this case is outlined.

The Chebyshev polynomial of degree  $n$  is

$$T_n(x) = \cos[n \cos^{-1}(x)] \quad (13)$$

Using elementary trigonometry, one can show the following recursion formula

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (14)$$

for  $n \geq 1$  with the initial condition  $T_0(x) = 1$ . The set of Chebyshev polynomials is a complete orthonormal set in the interval  $[-1,1]$  relative to the measure  $d\mu(x) = (1-x^2)^{-1/2}dx$ . The orthogonality property of the Chebyshev polynomials is given by the formula

$$\int_{-1}^1 \frac{T_i(x)T_j(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0 & i \neq j \\ \pi/2 & i = j \neq 0 \\ \pi & i = j = 0 \end{cases} \quad (15)$$

Note that this orthogonality condition is only for the measure  $d\mu(x)$ . The Chebyshev polynomials also satisfy the discrete orthogonality relation

$$\sum_{n=1}^N T_i(x_n)T_j(x_n) = \begin{cases} 0 & i \neq j \\ N/2 & i = j \neq 0 \\ N & i = j = 0 \end{cases} \quad (16)$$

where  $x_n$ s range over the zeros of  $T_M(x)$ . Note that the zeros of  $T_k$  are

$$x_n = \cos \frac{\pi(n - \frac{1}{2})}{k} \quad (17)$$

In Fig. 3, the Chebyshev polynomials of degree up to five are shown. Since the points  $\{x_n\}$  are not uniformly spaced and the planetary ephemerides are computed at equally spaced time intervals, the application of the Chebyshev polynomials is hardly appropriate. In the next section, the Gram polynomials are described, which are more suitable for this application.

## B. Gram Polynomial Equations

The Gram polynomial equations are most suitable for obtaining approximations to the planetary ephemerides or other data obtained by sampling at equally spaced time intervals.

The Gram polynomial [8]  $p_n(x, 2L)$  is defined by

$$p_n(x, 2L) = \sum_{k=0}^n (-1)^{k+n} \frac{(j+k)^{(2k)}}{(k!)^2} \frac{(L+x)^k}{2L^{(k)}} \quad (18)$$

where

$$x^{(n)} = x(x-1)(x-2)\cdots(x-n+1) = \prod_{j=0}^{n-1} (x-j)$$

with  $x^{(0)} = 1$ . Gram polynomials satisfy the orthogonality relations

$$\sum_{k=-L}^{k=L} p_i(k, 2L)p_j(k, 2L) = 0 \quad \text{for } i \neq j$$

$$\sum_{k=-L}^{k=L} p_i^2(k, 2L) = \frac{(2L+i+1)!(2L-i)!}{(2i+1)[(2L)!]^2} \quad (19)$$

Figure 4 gives the graphs of the polynomials  $p_1, \dots, p_5$ , for  $L = 10$ .

## V. Piecewise Polynomial Approximation

In this section, the result of Section III is extended to take into account the boundary conditions. The motivation for this extension is that the interpolated function approximating the planetary ephemerides is used as an update for the frequency control to the DFS, and it is necessary for this function to be continuous. It is shown later in Section VI that considerable improvement is achieved in the MSE, when the ephemerides data are subdivided into blocks, and each block is approximated using a different set of polynomials. To incorporate the boundary values, the values at the end-points are introduced as constraints in the original minimization problem, i.e., the following minimization problem is considered

$$\text{Min}_{\mathbf{a}} \chi^2(\mathbf{a}) = \int_a^b \left| f(x) - \sum_{i=1}^M a_i \phi_i(x) \right|^2 d\mu \quad (20)$$

subject to:

$$\sum a_i \phi_i(1) = A$$

$$\sum a_i \phi_i(-1) = B$$

Evaluating the partial derivative of  $\chi$  with respect to  $a_i$  and setting it equal to zero yields the following set of equations:

$$\begin{aligned} \frac{\partial \chi}{\partial a_j} = 0 &\Rightarrow \sum_k \left( \sum_i \phi_k(x_i) \phi_j(x_i) \mu(x_i) \right) a_k \\ &= \sum_i f(x_i) \phi_j(x_i) \mu(x_i) \end{aligned} \quad (21)$$

These equations form a system of  $M+2$  equations in  $M+2$  unknowns, namely  $Ca = L$  where

$$C = \left[ \begin{array}{cccc|cc} C_{ij} = \sum_k \phi_i(x_k)\phi_j(x_k)\mu(x_k) & & & & -\phi_1(1) & -\phi_1(-1) \\ & & & & -\phi_2(1) & -\phi_2(-1) \\ & & & & \vdots & \vdots \\ & & & & -\phi_M(1) & -\phi_M(-1) \\ \hline & \phi_1(1) & \phi_2(1) & \dots & \phi_M(1) & \\ & \phi_1(-1) & \phi_2(-1) & \dots & \phi_M(-1) & \\ & & & & 0 & 0 \\ & & & & 0 & 0 \end{array} \right] \quad (22)$$

and the vector  $\mathbf{L}$  is given by

$$\mathbf{L}_j = \sum_i f(x_i)\phi_j(x_i)\mu(x_i) \quad \text{for } 1 \leq i \leq N \quad (23a)$$

$$\mathbf{L}_{N+1} = A \quad \text{and} \quad \mathbf{L}_{N+2} = B \quad (23b)$$

In the next section, the performance of the algorithms introduced here will be evaluated by numerical simulations for a number of cases.

## VI. Numerical Simulations

The simulation results in this section are based on using one of the two following orthogonal bases. The Chebyshev polynomial equations of degree five or less in the interval  $[-1,1]$  are

$$\Delta = \{1, y, -1 + 2y^2, -3y + 4y^3, \\ 1 - 8y^2 + 8y^4, 5y - 20y^3 + 16y^5\} \quad (24)$$

The Gram polynomials of degree five or less in  $[-120,120]$  are

$$\Delta = \left\{ 1, \frac{x}{120}, \frac{121}{239} + \frac{x}{9560}, -\frac{43559}{3412920}x + \frac{x^3}{682584}, \right. \\ \left. \frac{7381}{18881} - \frac{12445}{46220688}x^2 + \frac{x^4}{46220688}, \right. \\ \left. \frac{37639643}{2272517160}x - \frac{9679}{1818013728}x^3 + \frac{x^5}{3030022880} \right\} \quad (25)$$

In the first set of experiments, the performance of each one of the proposed techniques for a discrete time function is compared. The original function  $y(x)$  is a sampled

second-order Chebyshev polynomial translated into the interval  $[0,1000]$ . This choice was intentionally made to show that even for a uniformly sampled Chebyshev polynomial, the Gram polynomial approximation outperforms the classical Chebyshev polynomial approximation. The results of this experiment are shown in Figs. 5 and 6. In Fig. 5 the original function is shown with 240 uniformly spaced samples between 0 and 1000. The interpolated function is not shown, since it is very close to the original function. The error sequence between the original sampled sequence and the approximating function resulting for each method is shown in Fig. 6. The corresponding MSE from Eq. (3) and the resulting polynomials are given in Table 1.

In Table 1, the MSE decreases by an order of one half when the number of sampled points in the original function is increased from 240 to 1000 points. The least squares method in this case was solved by using the SVD, and it gives a smaller MSE than the orthonormal expansion method. Note that the MSE is lower in each case when the Gram polynomials are used for approximating the original function, which in this case is itself a sampled third-order Chebyshev polynomial.

In the next two sections, these techniques are directly applied to the cases of this study, namely, the ephemerides data.

### A. Phobos Experiment

Phobos is a Martian moon. It completes an orbit of Mars approximately each eight hours, and its high speed accounts for one of the highest Doppler shifts encountered in the solar system. For this reason, Phobos' ephemerides were chosen for this case study. The original ephemerides are obtained at one-half-minute intervals, resulting in 960 points. Here, the orthonormal expansion methods, as described in Section III, are used. The method of least squares becomes prohibitively complex with 960 points.

In Fig. 7(a) the Doppler shift is shown. The resulting error sequence for each technique is shown in parts (b) and (d). The interpolated function is given in Fig. 7(c).

Note that the error sequence in Fig. 7(c) is very close to zero when shown in full scale. The fluctuation of the error is within 5 percent of the full scale of the Doppler shift. From Table 2, it is deduced that the constrained piecewise Gram polynomial approximation is superior, in terms of the MSE, to the orthonormal expansion.

## B. Comet Experiment

Another interesting experiment is based on the data simulating a celestial body (such as a comet or an asteroid) approaching the Earth at high speed. This is referred to as the comet experiment. The results of this test are shown in Fig. 8.

When the error sequences are compared, it becomes obvious that the piecewise approximation method reduces the end-point error by a factor of fifteen. From Table 3, it is seen that the constrained piecewise polynomial approximation outperforms the orthonormal expansion method.

It is concluded from the numerical simulation in this section that the constrained piecewise Gram polynomial approximation has the least MSE.

## VII. Frequency and Phase Error Due to the Update Rate

There are four sources of error: (a) phase error between the approximated Doppler and the actual Doppler, for the whole period of the ephemerides, (b) phase tracking error, which is the phase error between the PLO output phase and the actual phase of the ephemerides, (c) frequency error between the approximated Doppler frequency and the actual frequency, and (d) frequency tracking error, which is the frequency error between the PLO output frequency and the actual frequency of the ephemerides.

A key design parameter for using the PLO is the update rate. This update rate must be chosen such that the phase errors over the tracking period of the celestial body do not exceed 1.2 degrees and the frequency errors also are kept under 2 hertz.

The ephemerides phase and the frequency error are assessed by using the polynomials from the piecewise Gram polynomial approximation method, shown in Tables 2 and 3. Each polynomial is computed at the rate of once for each update period. The phase error between interpolated ephemerides and the actual ephemerides is evaluated by hard quantizing the interpolated function and the original function and computing the phase difference between each waveform. The frequency error is found by evaluating the

largest deviation between the interpolated function (computed at the update rate) and the original function, i.e.,

$$\Delta f = \text{Sup}_t | f(t) - \hat{f}(t) | \quad (26)$$

The results are shown in Fig. 9 for the Phobos and the comet experiments.

It follows from Fig. 9(a) that to maintain the phase errors under 1 degree during the whole tracking period, and the frequency errors under 2 hertz, the update period must be chosen to be less than 50 milliseconds.

The effects of the update rate on the frequency tracking error for the output spectrum of the PLO can be analyzed by considering a small segment (e.g., 1 minute) of the frequency error variation between the synthesized and the original function. During this period, this variation can be modeled as a ramp shown in Fig. 10(a). It should be noted that this is a valid approximation since the period of the ephemerides ( $\geq 5$  hours) is much larger than the update rate, which ranges between 50 milliseconds and 2 seconds (slowly moving celestial bodies). During each update rate, the frequency is either increased or decreased for a long period (usually in minutes).

The output frequency of the DFS (ranging between 10 to 20 MHz) is approximated by the linear function whose slope is

$$\frac{df}{dt} \approx \frac{\hat{f}(t) - \hat{f}(t - T)}{T} \quad (27)$$

and at the midpoint of the update period is equal to the value of the interpolating function. The resulting phase error is simply the integral of the frequency error, as shown in Fig. 10(a)

$$\Delta\phi(t) = 2\pi \int_0^t \Delta f(\tau) d\tau \quad (28)$$

In Fig. 10(b), the periodic phase error (in degrees) is shown as a function of the update period  $T$ . For example, in the case of the planets and their moons, one can roughly approximate their orbits as sinusoids [see Fig. 10(c)] and the resulting phase error is given in Fig. 10(d). Note that when the update period is very small compared to the ephemerides period, one can locally approximate the sinusoid by the ramp function. Then the phase error becomes approximately  $\Delta\phi = 2\pi(df/dt)T^2/8$ , see Fig. 10(b).

The effect of frequency tracking error can be approximated by using classical results from frequency modulation (FM) theory. Let  $p(t)$  be

$$p(t) = \begin{cases} \pi \left[ \left( \frac{t}{T} \right) - \frac{t}{T} \right]^2 & 0 < t \leq T \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

The output spectrum of the PLO is found by evaluating the spectrum of the signal  $o(t)$ , which is

$$o(t) = \sin \left( \omega_o t + \beta \sum_k p(t - kT) \right) \quad (30)$$

In FM modulation, the modulation index controls the spectral characteristics of the signal. In this case, the modulation index  $\beta = 2\pi(df/dt)T^2/8$ , where  $df/dt$  is the slope of the frequency ramp.

The pulse sequence in the phase term of Eq. (3) has the Fourier series expansion as

$$\sum_k p(t - kT) = -\frac{\pi}{6} + \frac{1}{\pi} \sum_{k=1}^{\infty} \left[ \frac{\cos(2\pi kt/T)}{k^2} \right] \quad (31)$$

Therefore,  $o(t)$  is a multi-tone FM signal for  $k$  ranging over a finite set. It can be shown, as in [10], that the spectrum of this signal can be approximated by certain sums of products of Bessel functions of the following form

$$\underbrace{\sum_{n_1} \sum_{n_2} \cdots \sum_{n_L}}_L \prod_{i=1}^L J_{n_i} \left( \frac{\beta}{i^2} \right) \cos \left[ \omega_c t + \sum_{k=1}^L \left( \frac{2\pi k n_k}{T} \right) t \right] \quad (32)$$

The series in Eq. (32) is composed of line spectrums in frequency domain, with a carrier component of amplitude

$J_0(\beta), \dots, J_0[\beta/(L-1)^2], J_0(\beta/L^2)$ . The analytical evaluation of the magnitude of each term in Eq. (32) is difficult, due to the intermodulation products. However, if the magnitudes of the harmonics of  $f_o - 2\pi/T$  and the intermodulation products are below  $-98$  dBc, then they are masked in the output spectrum by the quantization noise induced internally in the DFS (the spectral purity of the DFS is  $-98$  dBc [4]).

The authors have developed a program for the numerical computation of each term in Eq. (32). Here, an example for Phobos is presented, which represents the worst case in terms of the rate of change of the frequency. Let  $df/dt = 30$  Hz/sec, and  $T = 3$  msec, then  $\beta = 3.75 \times 10^{-4}$ . In Fig. 10, the line spectra of two cases with  $\beta = 1/\pi$  and  $\beta = 3.75 \times 10^{-4}$  are shown. In Figs. 11(a) and 11(b), the integer  $n$  represents the frequency  $f_o - 2n\pi/T$ . Note that for the update period of 3 msec ( $\beta = 3.75 \times 10^{-4}$ ) the magnitude of the harmonics at  $f_o - 2\pi/T = 10002094$  Hz (with  $f_c = 10$  MHz) is around  $-100$  dBc, and the spectral lines due to the intermodulation products are below  $-300$  dBc. It follows that an update rate of 1 KHz (for Phobos) is sufficient to guarantee that the effects of the frequency update rate have impact on the spectral purity of the PLO.

## VIII. Conclusion

An algorithm is described for piecewise orthonormal expansion in terms of Gram polynomials. This method outperforms other approaches in terms of the MSE by a factor of one hundred in some cases. This algorithm was applied to a number of cases for the ephemerides approximation. Using the piecewise Gram polynomial approximation algorithm, the programmable local oscillator can operate at a minimum frequency update rate of 2.94 kHz to maintain a minimum worst-case phase error of at most 1.0 degree when tracking a moon such as Phobos, which represents one of the worst cases of Doppler shift that may be encountered in the solar system.

## Acknowledgments

The authors appreciate the support and suggestions of Dr. Ray Jurgens, and also thank Steve Morris, David Hills, and Dr. Loris Robinett for their support of this work, and Dr. M. Standish and Dr. D. K. Yeomans of the Navigation Systems Section for providing the ephemerides data.

## References

- [1] F. R. Jurgens, "Earth-Based Radar Studies of Planetary Surfaces and Atmospheres," *IEEE Trans. Geo. Rem. Sen.*, vol. GE-20, no. 3, July 1982.
- [2] H. C. Rumsey, G. A. Moris, R. R. Green, and R. R. Goldstein, "A Radar Brightness and Altitude Image of a Portion of Venus," *Icarus*, vol. 23, pp. 1-7, 1974.
- [3] X X Newhall, "Numerical Representation of Planetary Ephemerides," *Celestial Mechanics*, vol. 45, 1989.
- [4] R. Sadr, E. Satorius, L. Robinett, and E. Olsen, *Digital Frequency Synthesizer for Radar Astronomy*, JPL Publication No. 90-32, Jet Propulsion Laboratory, Pasadena, California, August 1990.
- [5] B. Conroy and D. Le, "Multipurpose Exciter with Low Phase Noise," *TDA Progress Report 42-97*, vol. January-March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 169-174, May 15, 1989.
- [6] L. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, New Jersey: Prentice Hall, 1975.
- [7] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge, England: Cambridge University Press, 1988.
- [8] A. Ralstan, *A First Course in Numerical Analysis*, New York: McGraw-Hill, Inc., 1965.
- [9] M. Schwartz, W. R. Bennet, and S. Stein, *Communication Systems and Techniques*, New York: McGraw-Hill, Inc., 1966.

**Table 1. Results of the approximation algorithms**

Approximation algorithm	Mean-square error	Polynomial
240 points Chebyshev orthonormal expansion	0.0773128	$- 0.865531 + 0.0735729x - 0.000826432x^2 + 0.00000228612x^3$
1000 points Chebyshev orthonormal expansion	0.0359122	$- 937736 + 0.0178576x - 0.0000479041x^2 + 3.19042 \times 10^{-6}x^3$
240 points least squares Chebyshev polynomial	$3.456 \times 10^{-6}$	$- 1.0 + 0.0746888x - 0.000826432x^2 + 0.00000228612x^3$
240 points least squares Gram polynomial	$1.81487 \times 10^{-15}$	$- 1.0 + 0.0746888x - 0.000826432x^2 + 0.00000228612x^3$
240 points Gram orthonormal expansion	0.0308202	$- 8.878124 + 0.0673012x - 0.000703089x^2 + 0.00000153851x^3$

**Table 2. The Phobos experiment**

Approximation algorithm	Mean-square error	Polynomial
Gram polynomial orthonormal expansion	9358.31	$- 1.00546 \times 10^6 - 1.83.043x - 3.5313x^2 + 0.00924722x^3 - 5.57485 \times 10^{-6}x^4$
Constrained piecewise Gram polynomial approximation	8936.54	$- 1.01784 \times 10^6 - 346.157x - 2.91032x^2 + 0.0083382x^3 - 5.12973 \times 10^{-6}x^4$

**Table 3. The comet experiment**

Approximation algorithm	Mean-square error	Polynomial
Gram polynomial orthonormal expansion	800.046	$- 127756 - 167.347x - 0.038991x^2 + 91.2606 \times 10^{-6}x^3 - 3.41438 \times 10^{-8}x^4$
Constrained piecewise Gram polynomial approximation	235.818	$- 130512 - 193.964x - 0.112562x^2 + 16.7592 \times 10^{-6}x^3 - 6.05596 \times 10^{-6}x^4$

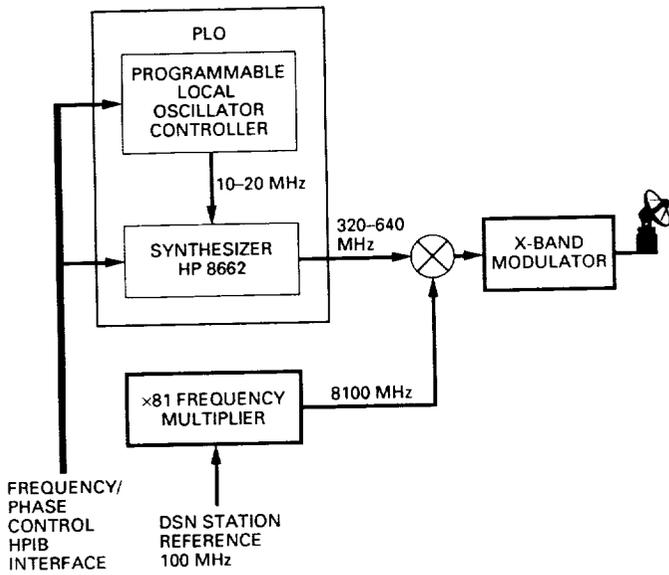


Fig. 1. X-band exciter for the Goldstone Solar System Radar Transmitter.

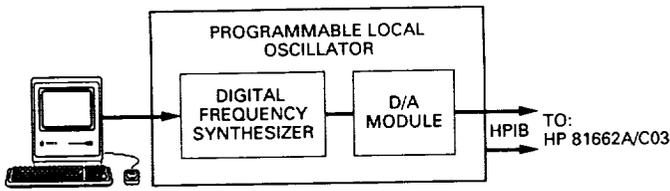


Fig. 2. Programmable local oscillator block diagram.

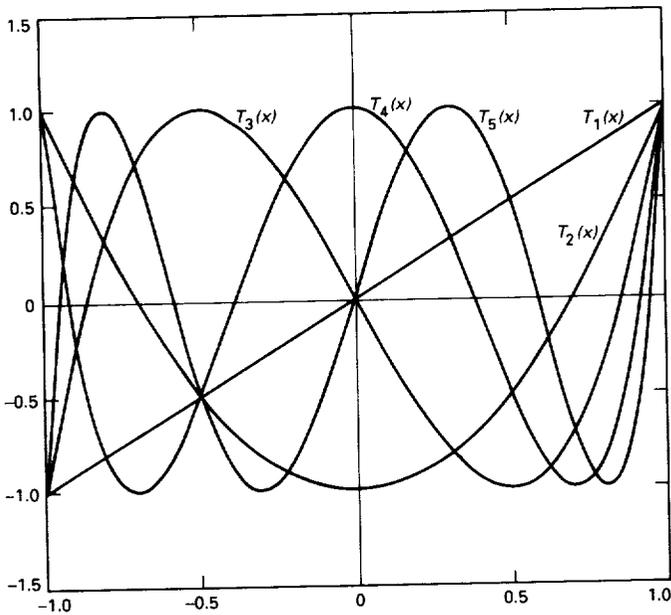


Fig. 3. Chebyshev polynomial equations.

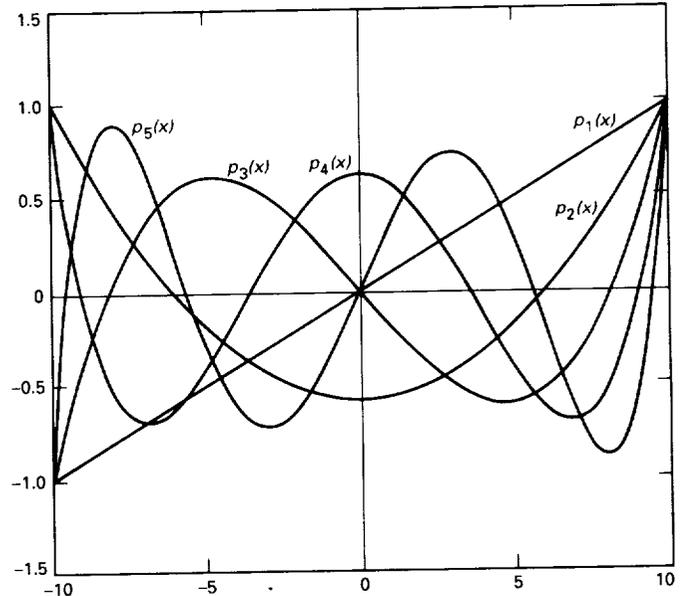


Fig. 4. Gram polynomial equations.

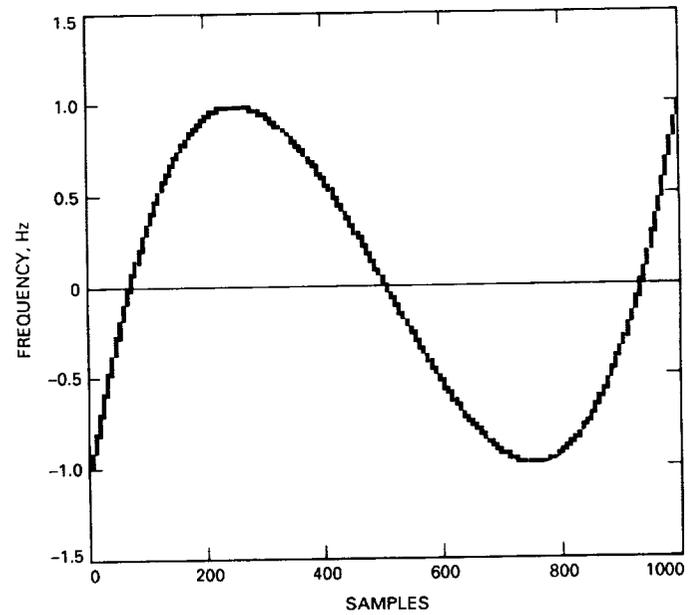


Fig. 5. Sampled original function using a third-order Chebyshev polynomial.

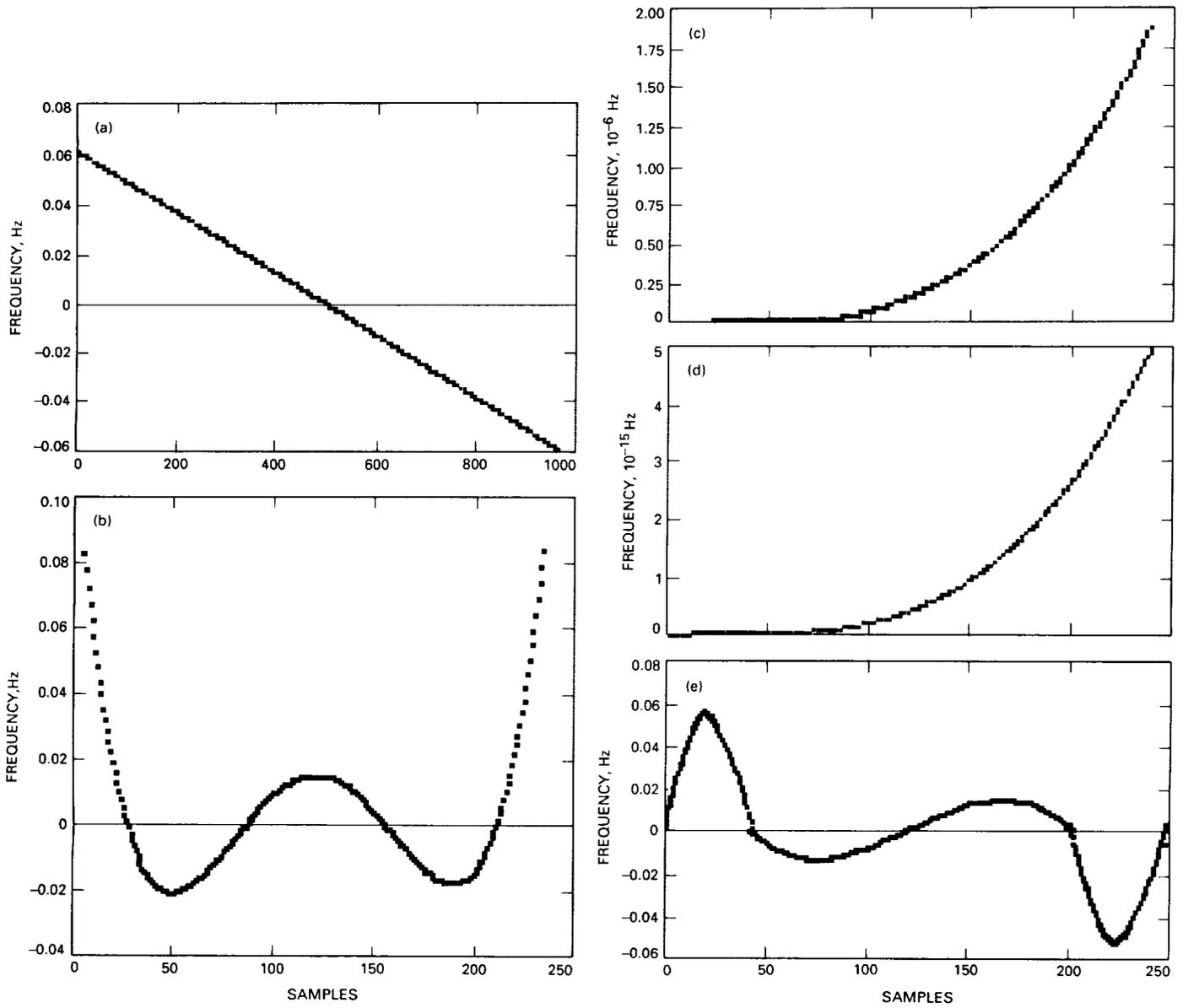


Fig. 6. Error function 240 points: (a) Chebyshev orthonormal expansion, (b) Gram orthonormal expansion, (c) Chebyshev least squares, (d) Gram least squares, and (e) Gram constrained piecewise polynomial approximation.

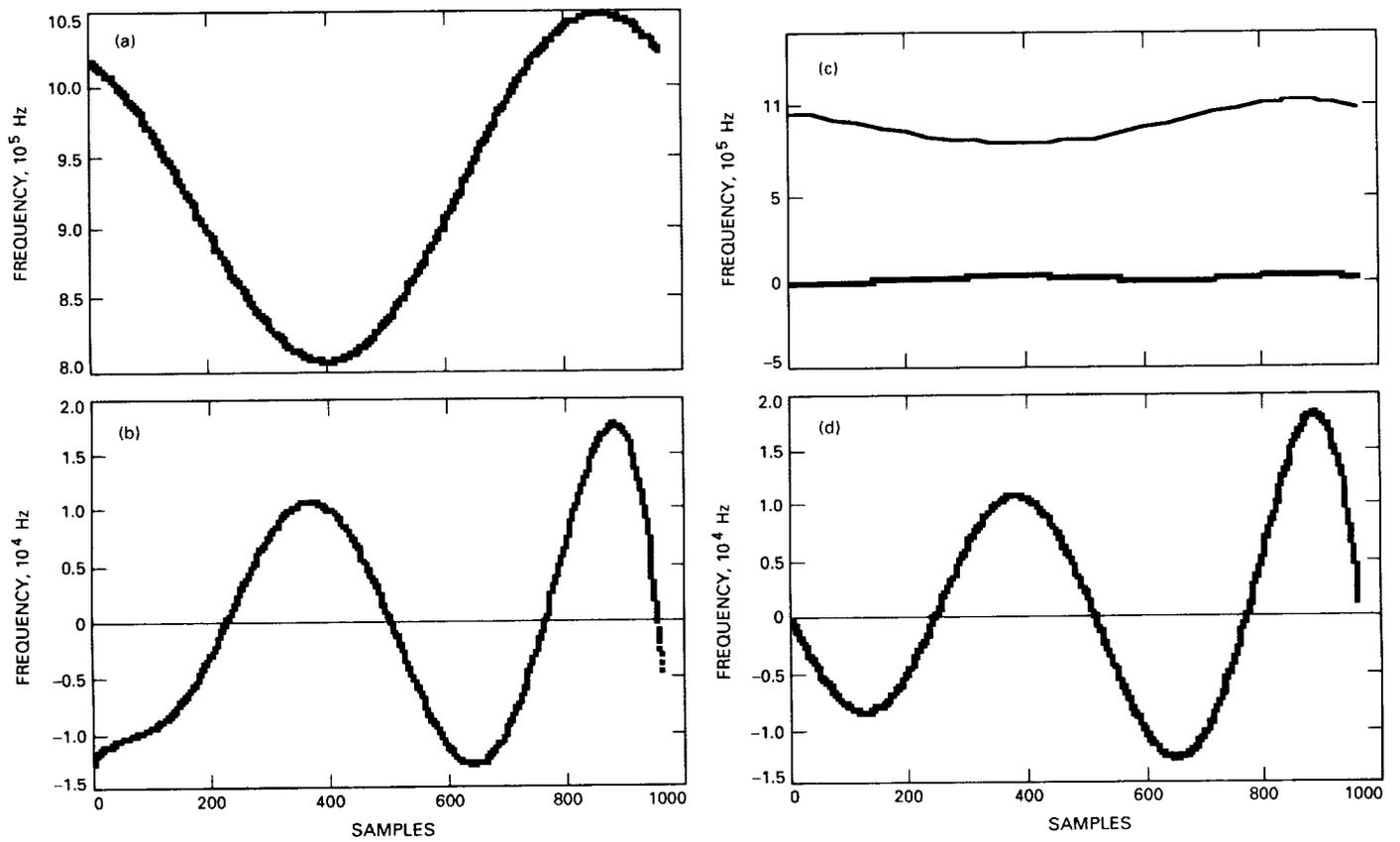
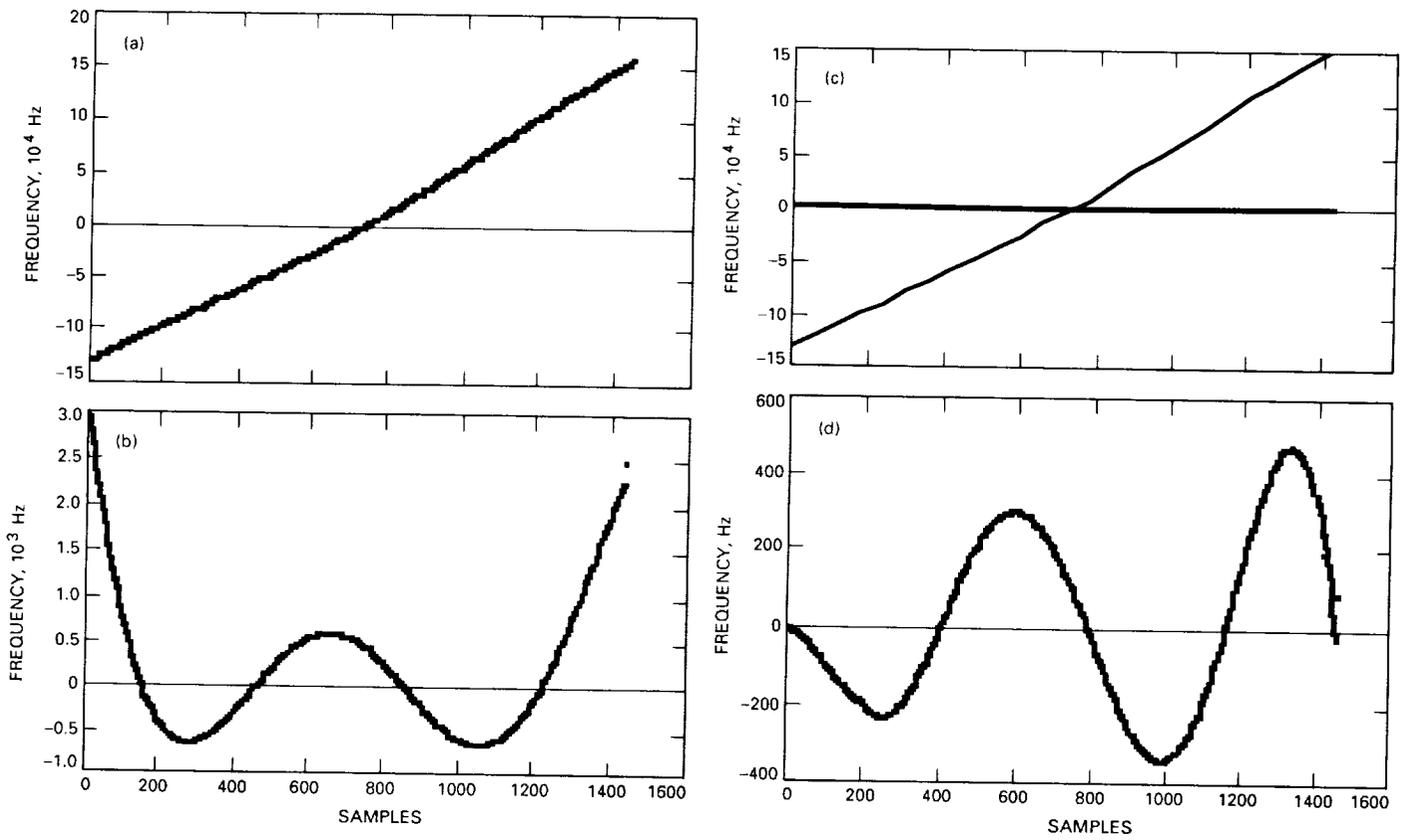


Fig. 7. The Phobos experiment: (a) Doppler shift, (b) error sequence, (c) error sequence and interpolated function, and (d) error sequence piecewise approximation.



**Fig. 8. The comet experiment: (a) Doppler shift, (b) error sequence, (c) error sequence and interpolated function, and (d) error sequence piecewise approximation.**

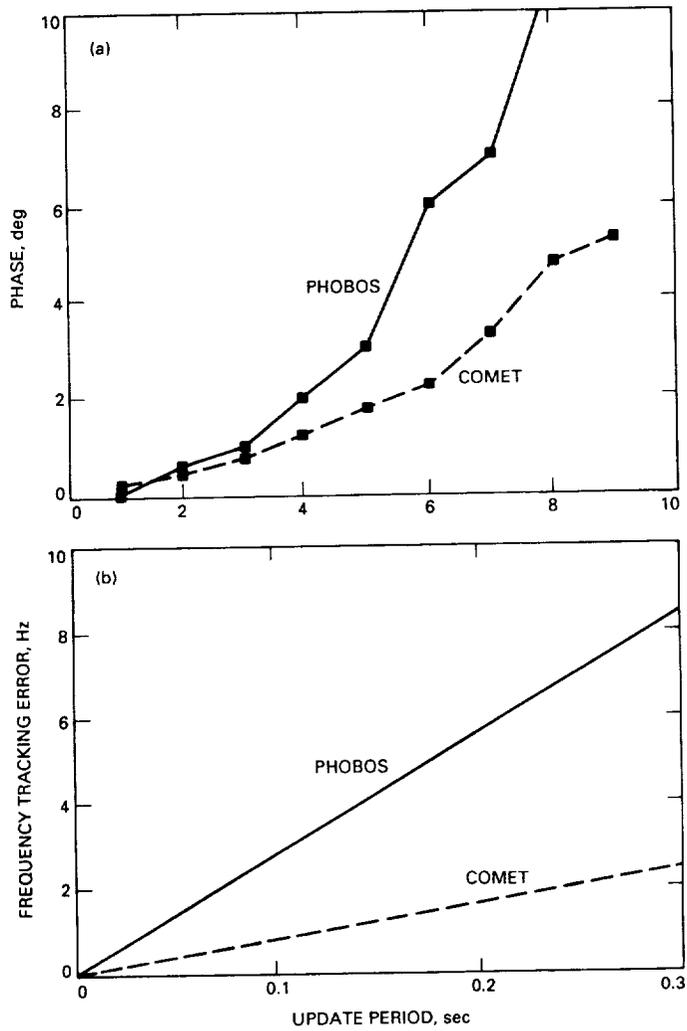


Fig. 9. Ephemerides tracking error: (a) phase versus update period and (b) frequency tracking error versus update period.

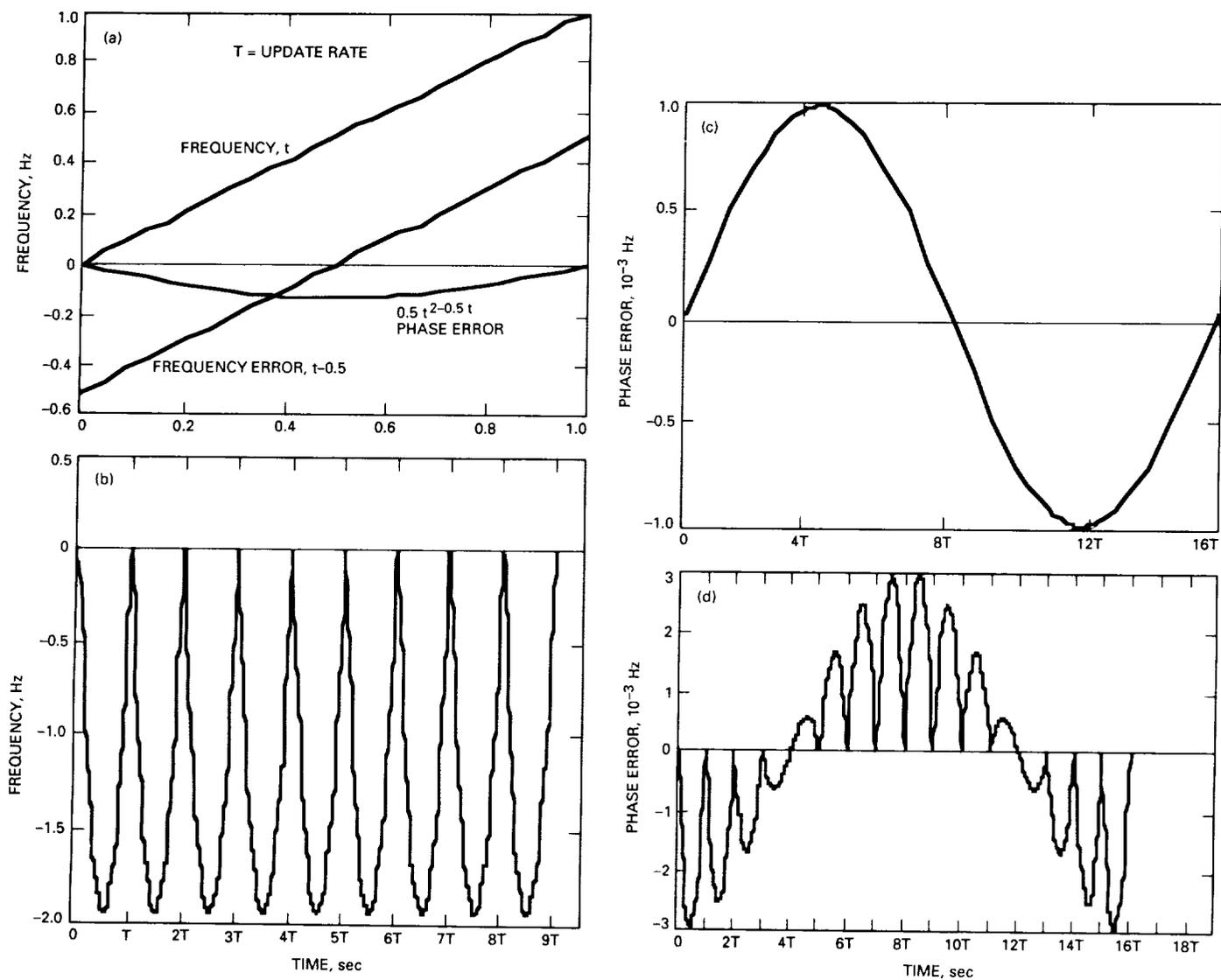


Fig. 10. Frequency and phase error due to Doppler: (a) frequency ramp and the phase error, (b) periodic phase error, (c) sinusoidal frequency ramp, and (d) phase error.

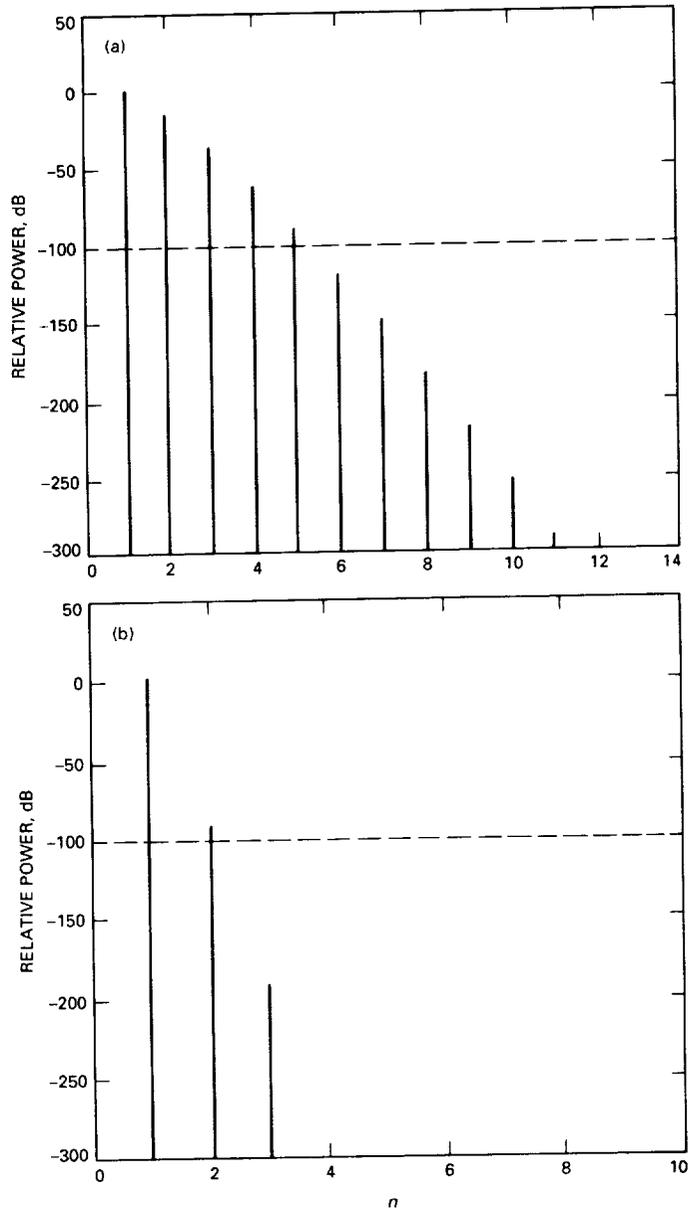


Fig. 11. Power spectra of an FM signal: (a)  $\beta = 1\pi$  and (b)  $\beta = 3.75 \times 10^{-4}$ .

N91-18327  
JJG 1990

# New Space Research Frequency Band Proposals in the 20- to 40.5-GHz Range

D. F. Bishop  
Telecommunications Systems Section

*Future space research communications systems may require spectra above 20 GHz. This article identifies frequency bands above 20 GHz that are suitable for space research. The selection of the proper bands depends on consideration of interference with other radio services, adequate bandwidths, link performance, and technical requirements for practical implementation.*

## I. Introduction

One of the areas of study in preparation for the World Administrative Radio Conference (WARC) of 1992 is the identification of bands above 20 GHz that could be used for space research. Figure 1 shows the existing allocations for space research (active) between 20 and 40.5 GHz. From Fig. 1, it can be seen that there are limited bands for space research. Therefore, a study was undertaken to identify additional bands between 20.0 and 40.5 GHz that are appropriate for space research. The criteria for selection of the new space research bands are provided. Several proposed frequency bands are chosen for consideration. Detailed band sharing studies between space research and existing services in these proposed bands are available.<sup>1</sup> Hardware considerations are discussed. Foot-

note 1 contains diagrams of hardware circuitry that can be used to provide coherence between the proposed space research bands and the existing space research allocations at 2 GHz (S-band) and 7 to 8 GHz (X-band).

## II. Criteria for New Space Research Bands

Several criteria went into the selection of potential space research bands. Minimum bandwidths of 500 MHz were sought for all proposed bands. Only bands between 20 GHz and 40.5 GHz were considered. Passive sensor bands in the space research services and radio-astronomy services were avoided to protect their sensitive receivers. To use the same antenna for transmit and receive operation, a minimum frequency spacing of 7 percent was desired. Coherence between the new band proposals and existing space research allocations at 2 GHz (S-band) and 7 to 8 GHz (X-band) was desired. Minimal interference with existing services in the proposed bands was sought. Since even small rainfall rates cause significant link degra-

<sup>1</sup>D. F. Bishop, "Background Material for Selection of New Space Research Frequency Bands Above 20 GHz," JPL Interoffice Memorandum 3396-90-90 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 26, 1990.

dation at frequencies above 20 GHz, it was not considered practical to select bands above 20 GHz based on weather conditions. Therefore, the effects of rain and clouds were not considered. The band selection in this article is based on clear-weather performance.

### III. Proposed Space Research Allocations

Based on the criteria of Section II, preliminary bands were selected as candidates for space research. Figures 2 and 3 show these bands. The band edges were selected to conform to the band edges that appear in the International Telecommunication Union (ITU) Table of Frequency Allocations ([1], page RR8-9). This was done to provide simple inclusion of a space research allocation should the proposal be accepted. Parts of these bands may not be suitable for space research due to hardware limitations or harmonic interference as demonstrated in the next section.

### IV. Radio Frequency Interference Analysis

One of the criteria for the selection of new space research bands was to have minimal interference with existing services. The potential interference was shown by performing sharing analyses between existing services and space research in the bands shown in Figs. 2 and 3. Two different scenarios were used for the space research services. In the first scenario, the space research spacecraft was located at the  $L_1$  Lagrangian (libration) point in the Earth-moon system [2]. Figure 4 contains an illustration of this scenario. In the second scenario, the space research spacecraft was located in a circular orbit around the Earth with an altitude of 500 km (Fig. 5). This is a typical orbit for a low Earth-orbiting spacecraft in the space research service.

Table 1 contains an example of a detailed sharing analysis. In this scenario, the 21.4- to 22.0-GHz band is proposed for space research in the space-to-Earth direction. In this frequency band there is a primary allocation for fixed and mobile services. Therefore, there is a potential for interference with a ground station in the space research service. Radio Regulation 2505 ([1], page RR27-2) permits a maximum Effective Isotropic Radiated Power (EIRP) of 55 dBW for a fixed or mobile service. This is the first entry on Table 1. The receiver bandwidth for a fixed radio relay system near 22 GHz is 40 MHz [3]. For simplicity, it is assumed that the total transmitter power is spread uniformly over this 40-MHz bandwidth; therefore, the spectral density level is  $-76.0$  dB/Hz ( $10 \cdot \log\{[40 \cdot 10^6]^{-1}\}$ ). This is the second entry on Table 1.

The third entry on Table 1 is the great-circle path loss. The great-circle path loss is selected to yield an interference margin of 0 dB in Table 1. The great-circle path loss determines a coordination distance [4]. A horizon angle of zero degrees is used. Larger horizon angles would yield lower interference levels at the Earth station receiver. Radio zone A2 is used. This zone contains Goldstone, California. The rate of attenuation due to all effects except atmospheric gases is determined for 0.001 percent of the time. This is the percentage of time that is recommended to protect deep-space research missions [5]. The result is that a coordination distance of 202 km between the fixed service transmitter and the space research Earth station is required.

The interference signal from the fixed service transmitter approaches the space research Earth station from the horizon. If it is assumed that the space research Earth station has an elevation angle of 10 degrees, then the gain of its antenna toward the horizon is 7 dBi ( $32 - 25 \cdot \log[10]$ ) from the ITU reference pattern [6]. This is the fourth entry on Table 1. The first four entries on Table 1 are added to yield the interference level at the space research Earth station receiver. The space research Earth station interference criterion is set equal to the deep-space Earth station interference criterion [7] at 32 GHz since deep-space Earth station receivers are often used for near-Earth space research, and it is closest in frequency to this proposed band from 21.4 to 22.0 GHz. The difference between the interference criterion and the interference level is the interference margin. In this case, the interference margin is 0.0 dB. Recall that the great-circle path loss was selected to yield an interference margin of 0 dB. This is the desired value. Coordination distances of greater than 202 km would yield positive values of interference margin. Therefore, the fixed services can share the 21.4- to 22.0-GHz band with space research (space-to-Earth) if the coordination distance between the stations is greater than 202 km measured over a great-circle path.

Footnote 1 contains the detailed sharing analyses that were performed in all of the proposed bands for space research. A summary of these sharing analyses is contained in Tables 4 through 12. These tables indicate the proposed frequency band, existing services in that band, the source of interference, the interference victim, the coordination distance for terrestrial interference scenarios, and the interference margin.

During the course of the study, another constraint on proposed bands was added. It was suggested that avoidance of harmonic interference from current space research allocations at 2 GHz (S-band) and 7 to 8 GHz (X-band)

is necessary. Therefore, the harmonics of the 2-GHz (S-band) and 7- to 8-GHz (X-band) allocations for space research that exist above 20 GHz were investigated. The allocations for near-Earth and deep-space research were considered because deep-space research equipment is often used for near-Earth space research. Table 2 contains some of the harmonic frequencies for space research (Earth-to-space) bands. The tenth harmonic of the space research allocation at 2 GHz (S-band) that is shown in Table 2 is probably too small to interfere with services in the harmonic band. Higher order harmonics should be even lower. The harmonics of the space research allocation at 7 to 8 GHz (X-band) that are shown in Table 2 may be large enough to cause interference. Therefore, these harmonic bands should be avoided by space research services. Space research ground stations that receive at these frequencies would be especially vulnerable to uplink harmonic emissions.

Table 3 contains some of the harmonic frequencies for space research (space-to-Earth) bands. The ninth harmonic of the space research allocation at 2 GHz (S-band) that is shown in Table 3 is probably too small to interfere with services in the harmonic band. Higher order harmonics should be even lower. The harmonics of the space research allocation at 7 to 8 GHz (X-band) that are shown in Table 3 may be large enough to cause interference. Therefore, these harmonic bands should be avoided by space research services. Space research spacecraft that receive at these frequencies would be especially vulnerable to these downlink harmonic emissions.

## V. Hardware Considerations

Hardware characteristics influence the selection of new bands for space research. Hardware design personnel indicated that a minimum spacing of 7 to 8 percent between the uplink and downlink frequencies is required to build practical diplexers. The use of a diplexer allows the same antenna to be used for uplink and downlink. The diplexer restricts the maximum spacing between uplink and downlink frequencies to 50 percent. Typically, space research services use circularly polarized signals. These circular polarizers limit the maximum spacing between uplink and downlink frequencies to 20 percent. Therefore, if a diplexer and a circular polarizer are used, the frequency spacing between uplink and downlink should be between 8 and 20 percent.

An additional hardware requirement is that the new space research allocations must be coherent with the existing allocations at 2 GHz (S-band) and 7 to 8 GHz

(X-band). The requirement is that the 2-GHz (S-band) downlink, the 7- to 8-GHz (X-band) downlink, and the new downlink (above 20 GHz) must be derived by the spacecraft transponder from the new uplink (above 20 GHz). Figure 6 contains an example of a circuit<sup>2</sup> that can provide coherence between the downlink frequencies and the uplink frequency. The uplink frequency of 24.46 GHz is within the proposed band of 24.25 to 25.25 GHz for space research (Earth-to-space). The circuitry derives an output frequency,  $F_a$  ( $4 \cdot 20.18$  MHz). This frequency is used as an input for three multiplier chains that derive the three downlink frequencies. The first downlink frequency of 27.77 GHz is within the proposed band of 27.5 to 29.5 GHz for space research (space-to-Earth). The second downlink frequency of 8.48 GHz is within the 8.45- to 8.50-GHz band for space research (space-to-Earth). The third downlink frequency of 2.26 GHz is within the 2.20- to 2.29-GHz band for space research (space-to-Earth). Therefore, coherence between the downlinks and the uplink has been provided. Additional circuits that provide coherence for spacecraft transponders are contained in Footnote 1.

## VI. Preferred Bands for Space Research

A comparison of link performance for the proposed space research bands in Figs. 2 and 3 was conducted in [8]. The conclusion was that the bands above 30 GHz provided improved link performance compared to the bands between 20 and 30 GHz. The downlink had up to 1-dB improvement and the uplink had up to 4.7-dB improvement. Above 30 GHz, the 30.0- to 31.3-GHz band (downlink) would be paired with the 34.2- to 35.2-GHz band (uplink) and the 37.0- to 37.5-GHz band (downlink) would be paired with the 39.5- to 40.5-GHz band (uplink).

Interference with other radio services in the 30.0- to 31.3-GHz band is minimal. The worst case interference occurs to a standard frequency and time signal satellite (SFTSS) Earth station from a low Earth-orbit space research satellite. In this case, there could be up to 26.5 seconds of interference during the worst orbit. All interference scenarios have low probabilities of occurrence.

In the 34.2- to 35.2-GHz band, there are some cases where interference can occur. Coordination of airborne-mobile or airborne-radiolocation stations in sight of space research stations is not practicable. Space research spacecraft within sight of terrestrial radiolocation stations

---

<sup>2</sup> Transponder circuits provided by C. E. Johns of the Radio Frequency and Microwave Subsystems Section, Jet Propulsion Laboratory, Pasadena, California.

would receive severe interference. Coordination of frequency assignments would be a method to reduce interference. In the 37.0- to 37.5- and 39.5- to 40.5-GHz bands, coordination distances are small enough that sharing between space research and other services is possible. Coordination with airborne stations is not practicable. Otherwise, the interference potential is minimal.

## VII. Summary and Conclusions

A search of the 20.0- to 40.5-GHz band has been made to identify frequency bands that may be used for space research. The criteria used for selection of the proposed bands are indicated. Several potential frequency bands are presented. Detailed band-sharing studies between space research and existing services in these proposed bands are provided in Footnote 1. Hardware constraints are discussed. Footnote 1 contains diagrams of hardware circuitry that can be used to provide coherence between the proposed space research bands and the existing space research allocations at 2 GHz (S-band) and 7 to 8 GHz (X-band). Lastly, preferred frequency bands for space research in the 20.0- to 40.5-GHz range are indicated.

Considering link performance and hardware implementation, the following band pairs are recommended for the space research service in the 20.0- to 40.5-GHz range:

30.0 to 31.3 GHz space research, space-to-Earth, primary in all three ITU regions, and

34.7 to 35.2 GHz space research, Earth-to-space, primary in all three ITU regions

or:

37.0 to 37.5 GHz space research, space-to-Earth, primary in all three ITU regions, and

39.5 to 40.5 GHz space research, Earth-to-space, primary in all three ITU regions.

From a hardware standpoint, the first band pair has the advantage of proximity to existing deep-space research allocations at 31.8 to 32.3 GHz and 34.2 to 34.7 GHz. However, the second band pair has a lower potential for interference with other radio services.

## References

- [1] *International Telecommunication Union Radio Regulations*, ISBN 92-61-01221-3, vol. 1, 1982.
- [2] J. Hall and J. Sos, "Preliminary Assessment and Implementation Plan for NASA Office of Exploration, Human Exploration Case Studies," MERTS Submission, *Space Frequency Coordination Group, Eighth Annual Meeting*, Buenos Aires, Argentina, 1988, Document SF/8-28, Section 3, p. 7, July 1988.
- [3] M. J. Pagonos, W. D. Rummmler, and W. J. Vidinski, "Protection of Terrestrial Line-of-Sight Radio-Relay Systems Against Interference from Space Stations in the Fixed-Satellite Service in Shared Frequency Bands Between 1 and 23 GHz," proposed revisions to *International Radio Consultative Committee (CCIR) Report 387-5*, CCIR Study Group 4/9, Document 01234567, CCIR, Geneva, Switzerland, p. 4, March 12, 1987.
- [4] CCIR, "Propagation Data Required for the Evaluation of Coordination Distance in the Frequency Range 1-40 GHz," *Recommendations and Reports of the CCIR, vol. V, Propagation in Non-Ionized Media*, CCIR Report 724-2, CCIR, Geneva, Switzerland, p. 485, 1986.
- [5] CCIR, "Protection Criteria and Sharing Considerations Relating to Deep-Space Research," CCIR Recommendation 578, *Recommendations and Reports of the CCIR, vol. II, Space Research and Radioastronomy*, CCIR, Geneva, Switzerland, p. 270, 1986.

- [6] *International Telecommunication Union Radio Regulations*, ISBN 92-61-01221-3, vol. 2, p. AP29-14, 1982.
- [7] CCIR, "Protection Criteria and Sharing Considerations Relating to Deep-Space Research," CCIR Report 685-2, *Recommendations and Reports of the CCIR, vol. II, Space Research and Radioastronomy*, CCIR, Geneva, Switzerland, Table IV, p. 282, 1986.
- [8] N. F. de Groot, *New Space Research Allocations in the 20-40 GHz Range*, United States Study Group for CCIR, Interim Working Party 2/2, draft report, CCIR, Geneva, Switzerland, p. 4, July 6, 1990.

**Table 1. Sharing analysis between fixed and mobile services and space research in the 21.4- to 22.0-GHz band**

Interference to a space research Earth station	
Fixed or mobile station EIRP (maximum allowed - ITU RR2505)	55.0 dBW
Spectral density level (uniform spectrum - 40-MHz bandwidth [3])	-76.0 dB/Hz
Great-circle path loss (202 km, $A_h = 0$ , zone A2, 0.001% of time)	-203.3 dB
Space research Earth station antenna gain toward the horizon (10-deg elevation angle, CCIR pattern)	7.0 dB
Interference level at space research Earth station receiver	-217.3 dBW/Hz
Space research interference criterion	-217.3 dBW/Hz
Interference margin	0.0 dB

**Table 2. Harmonic bands for space research (Earth-to-space)**

Space research allocation band, GHz	Harmonic number	Harmonic band, GHz
2.025 - 2.120	10	20.25 - 21.20
7.145 - 7.235	3	21.435 - 21.705
7.145 - 7.235	4	28.58 - 28.94
7.145 - 7.235	5	35.72 - 36.18
7.145 - 7.235	6	42.87 - 43.41

**Table 3. Harmonic bands for space research (space-to-Earth)**

Space research allocation band, GHz	Harmonic number	Harmonic band, GHz
2.20 - 2.30	9	19.80 - 20.70
8.40 - 8.50	3	25.20 - 25.50
8.40 - 8.50	4	33.60 - 34.00
8.40 - 8.50	5	42.00 - 42.50

**Table 4. Summary of sharing analysis for 21.4 to 22.0 GHz for space research (space-to-Earth)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Fixed and mobile	Fixed or mobile	Space research Earth station	202	0.0
	Space research spacecraft, Case 1	Fixed or mobile	—	24.8
	Space research spacecraft, Case 2	Fixed or mobile	—	-12.4

**Table 5. Summary of sharing analysis for 23.0 to 23.6 GHz for space research (Earth-to-space)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Fixed and mobile	Fixed or mobile	Space research spacecraft, Case 1	—	8.7
	Fixed or mobile	Space research spacecraft, Case 2	—	-28.5
	Space research Earth station	Fixed or mobile	250.4	0.0
Intersatellite	Relay satellite	Space research spacecraft, Case 1	—	-2.3
	Relay satellite	Space research spacecraft, Case 2	—	-22.2
	Space research Earth station	Intersatellite satellite	—	-113.4

**Table 6. Summary of sharing analysis for 24.25 to 25.25 GHz for space research (Earth-to-space)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Radio navigation	Radio navigation	Space research spacecraft, Case 1	—	-6.3
	Radio navigation	Space research spacecraft, Case 2	—	-31.9
	Space research Earth station	Radio navigation	331	0.0

**Table 7. Summary of sharing analysis for 25.25 to 27.0 GHz for space research (space-to-Earth)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Data relay satellite (candidate band)	Space research spacecraft	Data relay satellite	—	-11.5
	User (low Earth orbit) of data relay satellite	Space research Earth station	—	-63.6

**Table 8. Summary of sharing analysis for 27.5 to 29.5 GHz for space research (space-to-Earth)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Fixed satellite (E-S)	Fixed satellite Earth station	Space research Earth station	142.6	0.0
	Space research spacecraft, Case 1	Fixed satellite (geostationary)	—	17.0
	Space research spacecraft, Case 2	Fixed satellite (geostationary)	—	-1.4
Fixed and mobile	Fixed or mobile	Space research Earth station	196	0.0
	Space research spacecraft, Case 1	Fixed or mobile	—	24.8
	Space research spacecraft, Case 2	Fixed or mobile	—	-12.4

**Table 9. Summary of sharing analysis for 30.0 to 31.3 GHz for space research (space-to-Earth)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Fixed satellite (E-S), mobile satellite (E-S)	Fixed satellite Earth station	Space research Earth station	138.6	0.0
	Space research spacecraft, Case 1	Fixed satellite (geostationary)	—	17.0
	Space research spacecraft, Case 2	Fixed satellite (geostationary)	—	56.6
SFTSS <sup>a</sup> (S-E)	SFTSS satellite	Space research Earth station	—	-24.6
	Space research spacecraft, Case 1	SFTSS Earth station	—	8.9
	Space research spacecraft, Case 2	SFTSS Earth station	—	-47.4
Fixed and mobile	Fixed or mobile	Space research Earth station	192	0.0
	Space research spacecraft, Case 1	Fixed or mobile	—	24.8
	Space research spacecraft, Case 2	Fixed or mobile	—	-12.4

<sup>a</sup> Standard frequency and time signal satellite.

**Table 10. Summary of sharing analysis for 34.7 to 35.2 GHz for space research (Earth-to-space)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Fixed and mobile	Fixed or mobile	Space research spacecraft, Case 1	—	16.1
	Fixed or mobile	Space research spacecraft, Case 2	—	-21.1
	Space research Earth station	Fixed or mobile	273.1	0.0
Radiolocation	Radiolocation	Space research spacecraft, Case 1	—	-7.6
	Radiolocation	Space research spacecraft, Case 2	—	-63.9
	Space research Earth station	Radiolocation	342.5	0.0

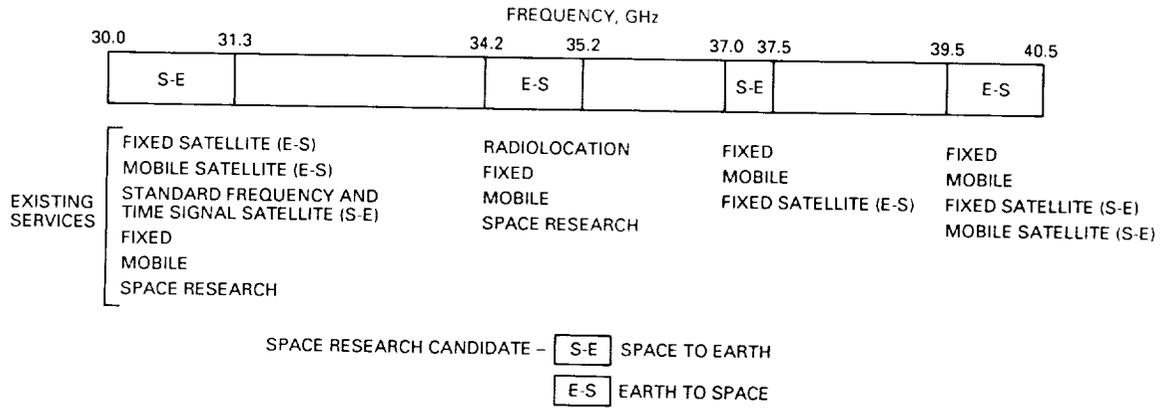
**Table 11. Summary of sharing analysis for 37.0 to 37.5 GHz for space research (space-to-Earth)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Fixed satellite (E-S)	Fixed satellite Earth station	Space research Earth station	120	0.0
	Space research spacecraft, Case 1	Fixed satellite (geostationary)	—	17.0
	Space research spacecraft, Case 2	Fixed satellite (geostationary)	—	58.3
Fixed and mobile	Fixed or mobile	Space research Earth station	168.5	0.0
	Space research spacecraft, Case 1	Fixed or mobile	—	24.8
	Space research spacecraft, Case 2	Fixed or mobile	—	3.5

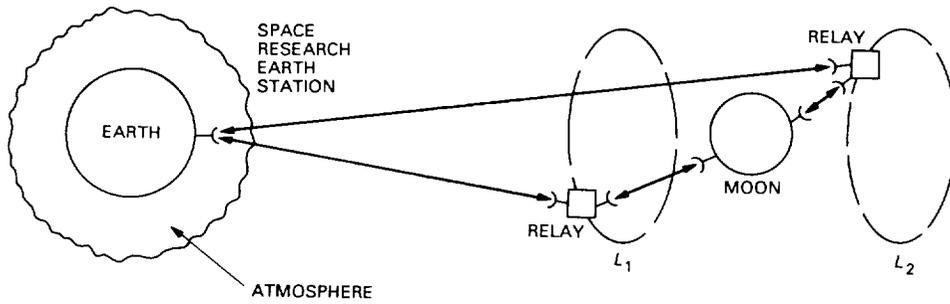
**Table 12. Summary of sharing analysis for 39.5 to 40.5 GHz for space research (Earth-to-space)**

Existing services	Interference source	Interference victim	Coordination distance, km	Interference margin, dB
Fixed satellite (S-E), mobile satellite (S-E)	Fixed satellite (geostationary)	Space research spacecraft, Case 1	—	11.2
	Fixed satellite (geostationary)	Space research spacecraft, Case 2	—	41.6
	Space research Earth station	Fixed satellite Earth station	198.8	0.0
Fixed and mobile	Fixed or mobile	Space research spacecraft, Case 1	—	16.1
	Fixed or mobile	Space research spacecraft, Case 2	—	7.2
	Space research Earth station	Fixed or mobile	238.2	0.0

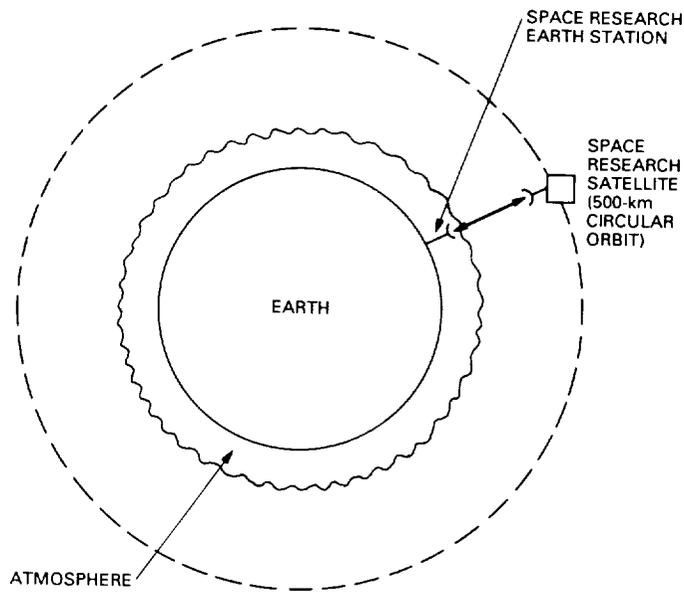




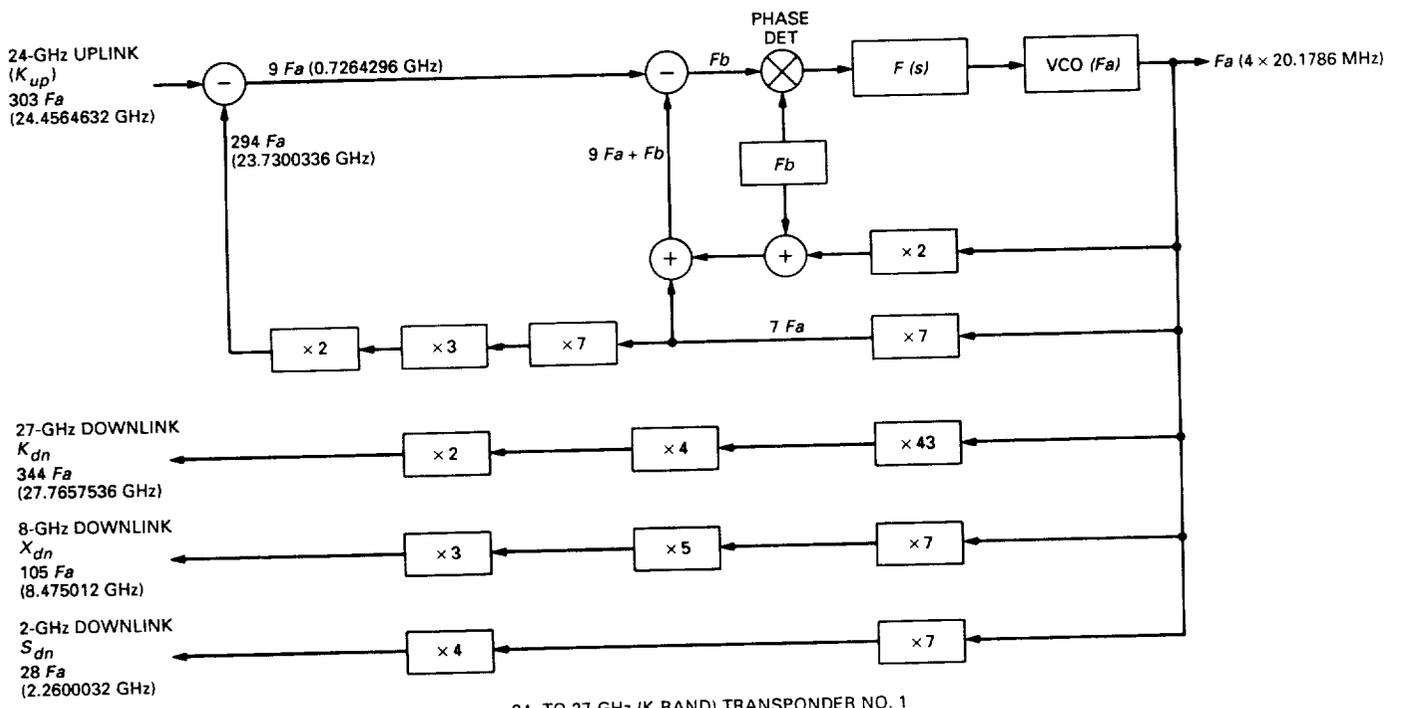
**Fig. 3. Candidate bands for space research sharing analysis, 30- to 40.5-GHz range.**



**Fig. 4. Space research spacecraft located near the Moon.**



**Fig. 5. Space research spacecraft located in low Earth orbit.**



FREQUENCY RATIOS:  
 24-GHz UPLINK/27-GHz DOWNLINK =  $303/344$     24-GHz UPLINK/2-GHz DOWNLINK =  $303/28$   
 24-GHz UPLINK/8-GHz DOWNLINK =  $303/105$     8-GHz DOWNLINK/2-GHz DOWNLINK =  $900/240$  ( $15/4$ )

Fig. 6. Example of a circuit used to provide coherence between the space research uplink and the downlinks.

520-32  
52075  
N91-18328 P-10

JJ574480

# Design Concepts and Performance of NASA X-Band (7162 MHz/8415 MHz) Transponder for Deep-Space Spacecraft Applications

N. R. Mysoor, J. D. Perret, and A. W. Kermode  
Spacecraft Telecommunications Equipment Section

*This article summarizes the design concepts and measured performance characteristics of an X-band (7162-MHz/8415-MHz) breadboard deep-space transponder (DST) for future spacecraft applications, with the first use scheduled for the Comet Rendezvous Asteroid Flyby (CRAF) and Cassini missions in 1995 and 1996, respectively. The DST consists of a double-conversion, superheterodyne, automatic phase-tracking receiver, and an X-band (8415-MHz) exciter to drive redundant downlink power amplifiers. The receiver acquires and coherently phase tracks the modulated or unmodulated X-band (7162-MHz) uplink carrier signal. The exciter phase modulates the X-band (8415-MHz) downlink signal with composite telemetry and ranging signals. The receiver measured tracking threshold, automatic gain control, static phase error, and phase jitter characteristics of the breadboard DST are in good agreement with the expected performance. The measured results show a receiver tracking threshold of  $-158$  dBm and a dynamic signal range of 88 dB.*

## I. Introduction

Telecommunication transponders for deep-space spacecraft applications provide independent uplink command and turnaround ranging functions, as well as downlink telemetry and radiometric capabilities. The spacecraft deep-space transponder (DST) is an element in the overall Deep Space Network (DSN) system. A balanced design approach for all the elements of the system must be applied to achieve end-to-end system performance capabilities that include telecommunications and radiometric

functions and multichannel and multifrequency capabilities [1]. The DST functions include:

- (1) precision phase/frequency reference transfer from the uplink signal
- (2) demodulation of the command and ranging signals from the uplink carrier
- (3) generation of a coherent or noncoherent downlink tracking signal for the Earth-based DSN

- (4) providing downlink signal modulation with composite telemetry data and turnaround ranging or differential one-way ranging (DOR) signals
- (5) providing a functional capability to utilize an external ultrastable oscillator (USO) to generate the downlink signal

This article describes the design, implementation, and performance of a breadboard DST configuration. The design specifications and functional description of the DST are summarized in Section II. The DST block diagram is described in Section III. The experimental results of a breadboard DST are presented in Section IV. Finally, some conclusions are drawn in Section V.

## II. Key Design Requirements

The design requirements for the DST are summarized in Table 1. The DST is to provide a receive and transmit capability at X-band (7162 MHz/8415 MHz) with the necessary reference signals to generate independent S-band (2295-MHz) and Ka-band (31,977-MHz) downlink signals external to the DST. Frequency translation ratios have been selected to provide coherent operation at S-band, X-band, and Ka-band, with overlap in all three frequency bands for simultaneous coherent operation. The selected transmit/receive frequency translation ratio for DST coherent operation at X-band (8415 MHz down, 7162 MHz up) is 880/749. The DST received uplink is at an assigned channel in the frequency range from 7145 to 7190 MHz ( $749 F_1$ ). The DST X-band ( $880 F_1 = 8415$  MHz) downlink frequency for the corresponding frequency channel assignment is in the frequency range from 8400 to 8450 MHz (Table 1). The receiver performance requirements include a maximum noise figure of 2.5 dB, a tracking threshold level of  $-158$  dBm, and a tracking range of  $\pm 250$  kHz at the assigned channel frequency. The acquisition and tracking rate is 550 Hz/sec at signal levels greater than  $-110$  dBm. The specified nominal output power of the exciter is  $+12.5$  dBm. The exciter output is phase modulated to a maximum phase deviation of  $\pm 2.5$  rad with a radio frequency (RF) modulation bandwidth greater than 40 MHz. The downlink phase noise requirements are 2.5 deg root mean square (rms) in the coherent mode and 2.8 deg rms in the noncoherent mode, when measured in a 10-Hz double-sided noise bandwidth DSN tracking receiver. The DST ranging and carrier phase delay variations over the flight acceptance (FA) temperature range ( $-10$  to  $+55$  deg C) are to be less than 22 nsec and 2.5 nsec, respectively. The differential downlink carrier phase delay variation is to be less than 1 nsec over the FA temperature range. The hardware qualification temperature range is from  $-20$  to

$+75$  deg C, with an expected flight operating range from  $+5$  to  $+45$  deg C. The hardware must withstand 15-krad (silicon) total radiation dose, 18.4-gravities (g) rms acceleration, 3000-g rms pyroshock, 12-g peak sine vibration, and 17.8-g rms random vibration environments.

## III. Transponder

### A. Block Diagram and Frequency Scheme

The DST frequency-generation scheme and functional block diagram are shown in Fig. 1. The receiver is implemented as a double-conversion, super-heterodyne, phase-lock tracking receiver, with a fixed-frequency second intermediate frequency (IF). The first local oscillator (LO) signal at  $880 F_1$  and the second LO signal at  $131 F_1 - F_2$  are generated by a dielectric resonator oscillator (DRO) [2,3] and a surface acoustic wave resonator oscillator (SRO), respectively. Both of these oscillators are phase locked to the 12  $F_1$  (114.75-MHz) voltage-controlled oscillator (VCO). The 12  $F_1$  VCO is in turn phase locked to the uplink carrier. The first and second intermediate frequencies are at  $131 F_1$  (1252.6875 MHz) and  $F_2$  (56.648 MHz), respectively. Coherent carrier automatic gain control (AGC) is employed in both of the IF sections to provide a constant signal plus noise at the carrier loop phase detector.

The coherent downlink carrier at  $880 F_1$  is provided by the LO DRO when the DST is operating in the coherent mode from the VCO. In the noncoherent mode, an  $880 F_1$  frequency is generated by the exciter DRO phase locked to the DST 12  $F_1$  auxiliary oscillator or the external USO. The noncoherent downlink signal is automatically selected by the receiver AGC function in the absence of an uplink signal. The DST's  $880 F_1$  phase-modulated signal [4] provides drive for the redundant spacecraft power amplifiers.

### B. Automatic Phase Tracking Loop

In the coherent mode, the  $880 F_1$  downlink signal generated in the DST exciter from the 12  $F_1$  VCO signal is phase coherent with the  $749 F_1$  received signal. Phase coherence is accomplished by an automatic phase-lock loop (PLL) in the receiver. The receiver PLL transfer function is a type-I, second-order lowpass filter [5,6]. The PLL design [5,6] is an involved iterative task and is usually a compromise between fast low-error tracking operation and noise response. The selection of the loop filter time constants ( $t_1$  and  $t_2$ ), the loop gain ( $K_V$ ), and the noise equivalent pre-detection bandwidth ( $B_L$ ), depends on six relevant receiver performance requirements. The requirements are:

- (1) the steady-state tracking error equal to 1 deg per 40 kHz offset at carrier levels greater than  $-110$  dBm

- (2) the minimum acceptable signal-to-noise power ratio (SNR) in the carrier channel at the phase detector input equal to  $-25$  dB
- (3) the minimum acquisition sweep rate at a strong signal ( $> -110$  dBm) equal to  $550$  Hz/sec
- (4) the damping factor at the theoretical threshold of ( $-158$  dBm) equal to  $0.5$
- (5) the damping factor at  $10$  dB above the theoretical threshold equal to  $0.8$
- (6) the two-sided loop noise bandwidth ( $2B_{LO}$ ) at the theoretical threshold ( $-158$  dBm) equal to  $18$  Hz

Using the above set of transponder performance requirements, the loop parameters  $t_1$ ,  $t_2$ , and  $K_V$  are selected using a PLL algorithm. Physical limitations of the components are also considered in this selection. Table 2 lists these loop parameters and compares them for several transponders: NASA Standard, Galileo, and Magellan.

### C. Residual Phase Noise

In the coherent carrier mode, residual phase noise is defined for a noise-free received signal case. Thus, phase noise on the downlink, unmodulated, carrier signal consists primarily of contributions from the three phase-locked oscillators  $12 F_1$  VCO, SRO, and DRO used in the DST implementation. Individual phase noise power spectral density functions [7,8,9] for these contributors are used in a comprehensive computer program to predict the phase noise of the closed-loop receiver. Total residual phase noise in the output is the mean square sum of all noise sources. The predicted phase noise for the DST in the coherent mode is shown in Fig. 2. In the intervals between  $5$  Hz and  $25$  MHz on each side of the carrier, the root mean square (rms) phase noise is  $0.448$  deg rms, which is well below the maximum allowable  $2.5$  deg for coherent downlink. The dominant contributor to this rms phase noise is the  $12 F_1$  VCO; the remaining contributions are less than  $10$  percent of the VCO contribution. Predicted rms phase noise and Allan deviation [8] are compared to the specified values in Table 3. The results of the analysis indicate that the coherent mode specifications will be met for both the rms phase noise and the Allan deviation. The closed-loop receiver servomechanism band limits the VCO spectrum, thus providing the superior performance in the coherent mode.

### D. Carrier Delay and Delay Variation

The phase variation associated with the temperature change of the transponder can be estimated by construct-

ing a model from the block diagram. The analysis assumes that the frequency multipliers are major contributors of the phase delay variation with temperature. The contribution for each multiplier is assumed to be three degrees of phase per degree Celsius. The estimated value of the DST carrier phase delay variations from input to output is equal to  $0.075$  nsec over the FA ( $-10$  deg C to  $+55$  deg C) temperature range. The predicted carrier delay data indicate that DST satisfies the requirement of maximum allowable delay variations equal to  $2.5$  nsec. Hardware performance characteristics over the temperature environment will be measured on an engineering model DST in 1991.

## IV. Transponder Experimental Results

A breadboard DST X-band receiver and exciter shown in Fig. 1 (without S-band and Ka-band exciters) was implemented and performance characterization accomplished in both the Transponder Development Laboratory and the Telecommunications Development Laboratory (TDL). The evaluation measurements include receiver tracking threshold sensitivity, static phase errors for X-band ( $7162$  MHz) uplink frequency offset, swept acquisition characteristics, and AGC versus uplink signal level. All measurements were made at room temperature ( $25$  deg C). The measured tracking threshold sensitivity at the receiver best lock frequency (BLF) (approximately channel center) is  $-158$  dBm, which is in good agreement with the design threshold value ( $-157.3$  dBm) using the receiver loop noise bandwidth of  $18$  Hz and the measured LNA noise figure of  $2.9$  dB at  $25$  deg C. The measured receiver threshold characteristics show good correlation with expected performance over the tracking range as shown in Fig. 3. The receiver acquisition characteristics were measured at an input signal level of  $-110$  dBm. The measured values for tracking range and tracking rate are  $\pm 270$  kHz at design center frequency and  $800$  Hz/sec, respectively, and meet the specified requirements (Table 1). Figure 4 shows a linear relationship for the static phase error voltage versus uplink frequency offset over the receiver tracking range. The AGC loop-filter amplifier-output voltage controls the gain in the first and second intermediate frequency (IF) amplifiers. The AGC voltage versus uplink signal level at the best lock frequency and at frequency offsets (FOs) of  $\pm 250$  kHz from BLF are shown in Fig. 5. As the receiver input signal varies from a strong signal level ( $-70$  dBm) to the threshold level ( $-158$  dBm), the AGC control voltage varies approximately linearly. No receiver false lock or self lock resulted during the test phase.

A comparison of measured to calculated phase jitter characteristics as a function of receiver uplink signal level

is shown in Fig. 6. The TDL-system-measured residual phase jitter at a downlink signal level of  $-100$  dBm with no uplink signal is equal to 2.62 deg rms. The TDL measured phase jitter values for the breadboard DST and Magellan transponders at the same uplink signal level of  $-100$  dBm are equal to 3.03 deg rms and 3.98 deg rms, respectively (referred to the same 2.62 deg rms TDL system residual phase jitter).

## V. Conclusions

Design concepts and system architecture for a high-performance X-band (7162-MHz/8415-MHz) DST for deep-space spacecraft applications have been presented.

The DST has been successfully breadboarded and evaluated. New technologies such as a dielectric resonator oscillator, X-band (8415-MHz) phase modulator, and SRO have been integrated into the design. The Telecommunications Development Laboratory measurements on the breadboard DST achieved a threshold level of  $-158$  dBm with a dynamic range of 88 dB and excellent acquisition and tracking characteristics. The measured tracking receiver threshold and phase jitter data are in good agreement with the predicted characteristics. The Jet Propulsion Laboratory breadboard X-band DST design and evaluation have demonstrated a basic model configuration for implementation of future deep-space transponders, and an engineering model development phase is expected to be completed by September 1991.

## Acknowledgments

The authors gratefully acknowledge the valuable support by John T. Meysenburg for breadboard fabrication and testing, Peter W. Kinman for phase delay calculations, and Keith Siemsen of MOTOROLA (Strategic Electronics Division, Chandler, Arizona) for phase noise calculations.

## References

- [1] J. H. Yuen, *Deep Space Telecommunication Engineering*, New York: Plenum Press, 1983.
- [2] N. R. Mysoor, "An Electronically Tuned, Stable 8415 MHz Dielectric Resonator FET Oscillator for Space Applications," *TDA Progress Report 42-93*, vol. January-March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 292-301, May 15, 1988.
- [3] N. R. Mysoor, "An Electronically Tuned, Stable 8415 MHz Dielectric Resonator FET Oscillator for Space Applications," *Proc. IEEE 1990 Aerospace Applications Conference*, Vail, Colorado, pp. 147-156, February 5-9, 1990.
- [4] N. R. Mysoor, "A Low-Loss Linear Analog Phase Modulator for 8415 MHz Transponder Applications," *TDA Progress Report 42-96*, vol. October-December 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 172-178, February 15, 1989.
- [5] F. M. Gardner, *Phaselock Techniques*, New York: John Wiley, 1979.
- [6] A. Blanchard, *Phase-Locked-Loops: Applications to Coherent Receiver Design*, New York: John Wiley and Sons, 1976.

- [7] D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proc. IEEE*, vol. 54, pp. 329–330, February 1966.
- [8] J. A. Barnes, A. R. Chi, and L. S. Cutler, *Characterization of Frequency Stability*, National Bureau of Standards Technical Note 394, Washington, D.C.: National Bureau of Standards, October 1970.
- [9] D. W. Allan, "Time and Frequency (Time-Domain) Characterization, Estimation, and Prediction of Precision Clocks and Oscillators," *IEEE Transactions on Ultrasonics, Ferroelectronics, and Frequency Control*, vol. UFFC-34, no. 6, pp. 647–654, November 1987.

**Table 1. Deep-space transponder specifications**

Parameter	Design requirement
1. Uplink frequency allocation	7145 to 7190 MHz, deep space
2. Downlink frequency allocation	8400 to 8450 MHz, deep space
3. Frequency translation ratios	
Channel 14 uplink frequency	7162.3125 MHz ( $749 F_1$ )
X-band downlink	880/749 (8415 MHz)
S-band downlink	240/749 (2295 MHz)
Ka-band downlink	3344/749 (31,977 MHz)
4. Receiver parameters	
Carrier threshold	-157.3 dBm
Dynamic range	88 dB (carrier threshold to -70 dBm)
Noise figure at DST receiver input	2.5 dB maximum over -20 deg C to +75 deg C
Acquisition and tracking rate	550 Hz/sec at signal level > -110 dBm
Tracking range	$\pm 250$ kHz minimum
Tracking error	< 1 deg/40 kHz at carrier level > -110 dBm
Capture range	$\pm 1.3$ kHz at signal level > -120 dBm
5. Exciter parameters	
Frequency for coherent operation	880/749 time uplink frequency
Frequency for noncoherent operation	$\approx 880F_1$
RF output power level	+12 dBm, nominal
Output impedance	$50 \pm 5$ ohms, nominal
Spurious signals	60 dBm below the carrier
Modulation bandwidth	> 40 MHz
Modulation index	Ranging: 3-9 dB carrier suppression Telemetry: 0-15 dB carrier suppression DOR: 0-1.1 dB carrier suppression
Modulation sensitivity	2 rad peak/volt peak
Modulation amplitude linearity	$\pm 2.5$ rad at $\pm 8$ percent linearity
Modulation index stability	$\pm 10$ percent over -20 deg C to +75 deg C
Residual phase noise	< 2.5 deg rms in the coherent mode < 2.8 deg rms in the noncoherent mode
Input-to-output carrier phase delay variation	< 2.5 nsec over -10 deg C to +55 deg C
Differential phase delay variation	< 1 nsec over -10 deg C to +55 deg C
Ranging phase delay variation	< 22 nsec over -10 deg C to +55 deg C

Table 2. Transponder carrier phase tracking loop parameters

Transponder	Design	Signal condition												
		Threshold, $\alpha_0 = 0.0531$					10 dB above threshold, $\alpha_{10} = 0.1665$					Strong signal, 50 dB above threshold, $\alpha_S = 1.0$		
		$K_v$ , $\text{sec}^{-1}$	$\tau_1$ , sec	$\tau_2$ , sec	$\zeta_0$	$\omega_{n_0}$ , rad/sec	$2B_{L_0}$ , Hz	$\zeta_{10}$	$\omega_{n_{10}}$ , rad/sec	$2B_{L_{10}}$ , Hz	$\zeta_S$	$\omega_{n_S}$ , rad/sec	$2B_{L_S}$ , Hz	$\Delta\omega$ , Hz/sec
NST		$1.44 \times 10^7$	2910	0.0833	0.68	16.23	17.0	1.20	40.3	28.7	2.90	70.3	209.9	394
NST + XSDC		$4.77 \times 10^7$	2910	0.0833	1.90	45.28	92.0	3.34	273.9	80.2	5.40	128.0	697.1	1314
GLL		$1.62 \times 10^7$	3732	0.0423	0.32	15.21	16.7	0.57	27.1	26.9	1.42	66.0	105.3	347
GLL + XSDC		$5.73 \times 10^7$	3732	0.0423	0.73	34.09	36.6	1.30	90.0	60.4	2.66	123.9	341.2	1222
MAG		$1.44 \times 10^7$	728	0.0208	0.42	40.38	41.0	1.49	118.3	71.5	1.50	126.0	210.0	1263
MAG + XSDC		$4.9 \times 10^7$	728	0.0208	1.00	75.00	120.0	1.38	207.3	132.8	2.70	228.0	636.7	4136
DST		$2.2 \times 10^7$	3556	0.0556	0.50	18.0	18.0	0.88	31.9	37.2	2.17	78.1	178.5	550

$K_v$  = DC gain of the PLL,  $\text{sec}^{-1}$

$\tau_1$  = Time constant associated with the open-loop pole (phase lag) of the loop filter, sec

$\tau_2$  = Time constant associated with the open-loop pole (phase lead) of the loop filter, sec

$\alpha_0$ ,  $\alpha_{10}$ ,  $\alpha_S$  = Limiter suppression factor at threshold, 10 dB above threshold, and strong signal

$\zeta_0$ ,  $\zeta_{10}$ ,  $\zeta_S$  = PLL damping factor at threshold, 10 dB above threshold, and strong signal

$2B_{L_0}$ ,  $2B_{L_{10}}$ ,  $2B_{L_S}$  = PLL noise-equivalent bandwidth at threshold, 10 dB above threshold, and strong signal, rad/sec

$\omega_{n_0}$ ,  $\omega_{n_{10}}$ ,  $\omega_{n_S}$  = PLL natural frequency at threshold, 10 dB above threshold, and strong signal, rad/sec

$\Delta\omega$  = PLL acquisition and tracking rate at strong signal, Hz/sec

NST = NASA Standard DST (S-band)

GLL = Galileo transponder (S-band)

MAG = Magellan transponder (S-band)

DST = Deep space transponder (X-band design)

XSDC = External X-band to S-band downconverter

**Table 3. Transponder coherent mode predicted rms phase noise and Allan deviation**

	Predicted output	Specification
RMS phase noise (deg rms)	0.448	2.5 (5 Hz to 25 MHz)
Allan deviation (integration time):		
0.01 sec	$2.6 \times 10^{-11}$	$3 \times 10^{-11}$
1.0 sec	$2.6 \times 10^{-13}$	$1.2 \times 10^{-12}$
1000 sec	$2.4 \times 10^{-15}$	$5 \times 10^{-15}$

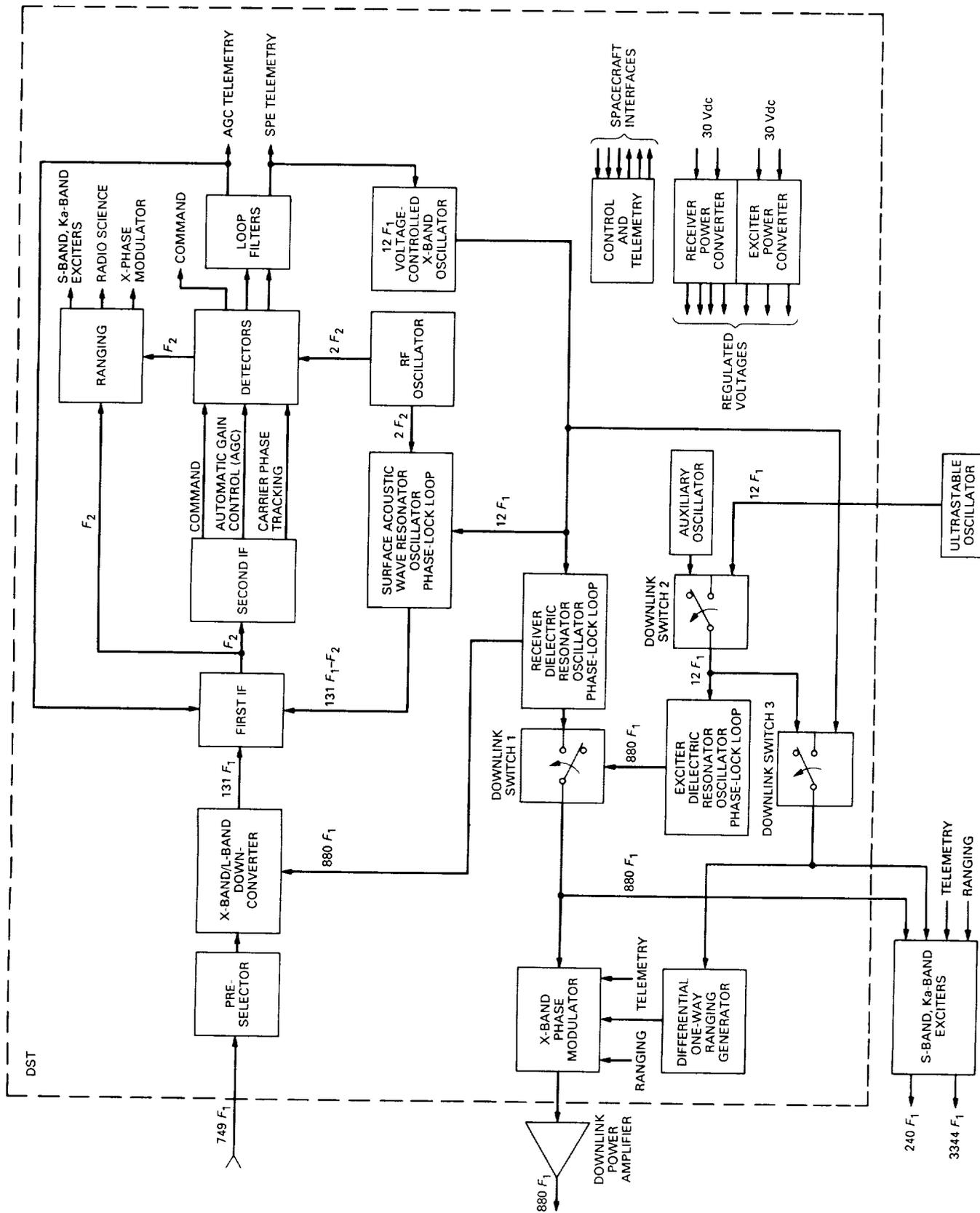


Fig. 1. Deep-space transponder.

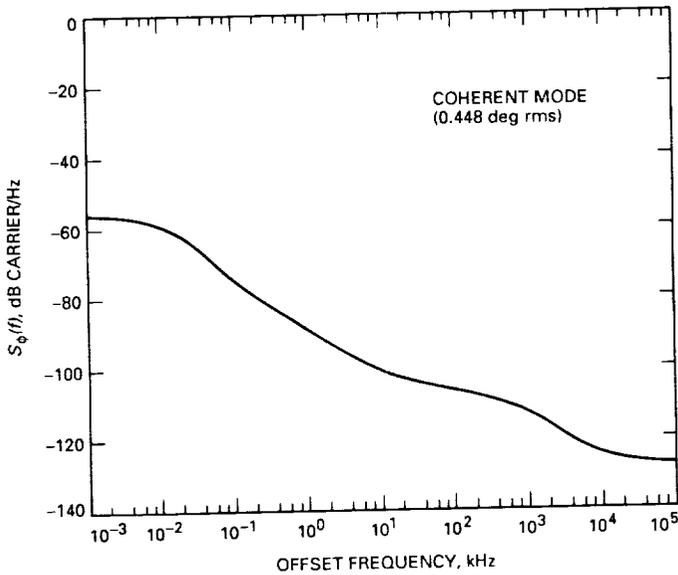


Fig. 2. DST 880  $F_1$  coherent mode phase noise.

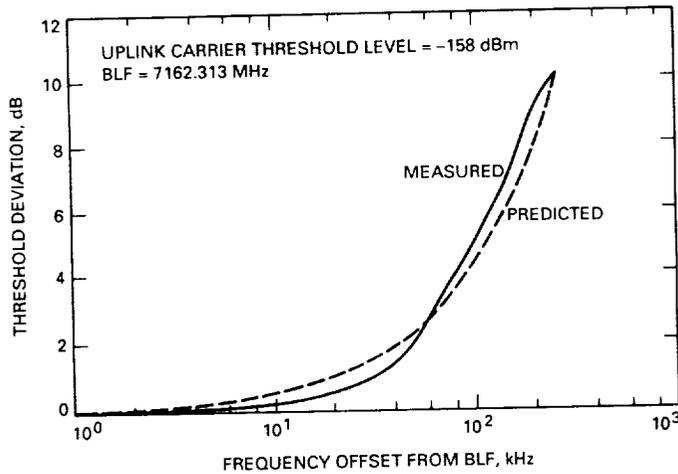


Fig. 3. DST carrier tracking threshold versus offset frequency.

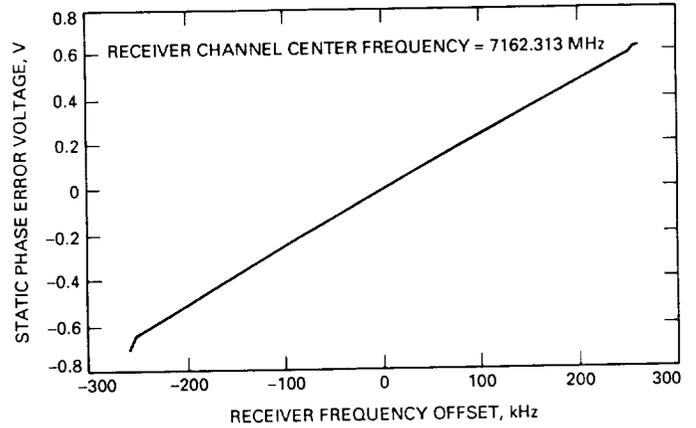


Fig. 4. DST static phase error voltage versus offset frequency.

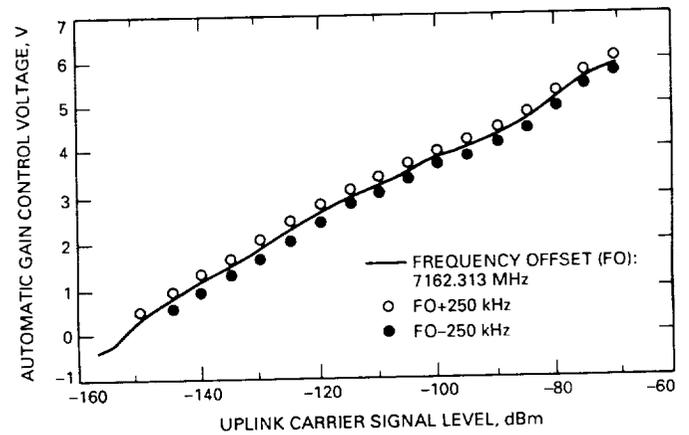


Fig. 5. DST AGC voltage versus uplink signal level.

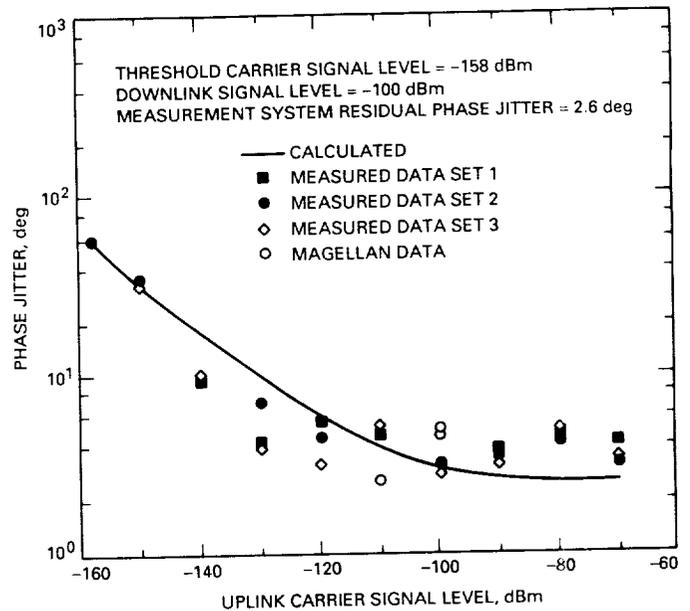


Fig. 6. DST phase jitter versus uplink signal level.

OMIT TO  
END

## Author Index,<sup>1</sup> 1990

### *The Telecommunications and Data Acquisition Progress Report*

*42-101, January–March, May 15, 1990*

*42-102, April–June, August 15, 1990*

*42-103, July–September, November 15, 1990*

*42-104, October–December, February 15, 1991*

#### Alvarez, L. S.

- 42-104 An Analysis of the Least-Squares Problem for the DSN Systematic Pointing Error Model, pp. 17–29.

#### Arnold, E. S.

- 42-104 Emerging Standards for Still Image Compression: A Software Implementation and Simulation Study, pp. 98–102.

See Pollara, F.

- 42-104 Frame Synchronization for the Galileo Code, pp. 211–218.

L. Swanson

- 42-104 Compressed/Reconstructed Test Images for CRAF/Cassini, pp. 88–97.

See Dolinar, S. J.

#### Bartos, K. P.

- 42-102 Interagency Telemetry Arraying for Voyager–Neptune Encounter, pp. 91–118.

See Brown, D. W.

#### Bassiri, S.

- 42-103 Three-Frequency Ranging Systems and Their Applications to Ionospheric Delay Calibration, pp. 14–20.

#### Batchelder, M. P.

- 42-103 Earth-Based Ranging Measurements to a Planetary Radio Beacon: An Approximate Analytic Formulation, pp. 70–86.

See Estefan, J. A.

#### Bathker, D.

- 42-101 Design and Performance Analysis of the DSS-13 Beam Waveguide Antenna, pp. 99–113

See Veruttipong, W.

#### Bevan, R.

- 42-104 The Advanced Receiver II: Telemetry Test Results in CTA 21, pp. 140–156.

See Hinedi, S.

#### Beyer, P. E.

- 42-103 Galileo Array Study Team Report, pp. 161–169.

See Layland, J. W.

<sup>1</sup> In the case of joint authorship, the reader is referred to the citation under the first author, where all the authors of the article are listed.

**Biraud, F.**

- 42-102 Technical Considerations on Using the Large Nançay Radio Telescope for SETI, pp. 152-160.  
See Gulkis, S.

**Bishop, D. F.**

- 42-104 New Space Research Frequency Band Proposals in the 20-40.5 GHz Range, pp. 235-246.

**Blasenheim, B. J.**

- 42-101 A Study of the Influence of Hg ( $6^3P_2$ ) Population in a Low-Pressure Discharge on Mercury Ion Emission at 194.2 nm, pp. 61-72.  
See Maleki, L.

**Brown, D. W.**

- 42-102 Interagency Telemetry Arraying for Voyager-Neptune Encounter, pp. 91-118.  
W. D. Brundage, J. S. Ulvestad, S. S. Kent, and K. P. Bartos
- 42-103 Galileo Array Study Team Report, pp. 161-169.  
See Layland, J. W.

**Brundage, W. D.**

- 42-102 Interagency Telemetry Arraying for Voyager-Neptune Encounter, pp. 91-118.  
See Brown, D. W.

**Brunzie, T. J.**

- 42-102 The Parkes Front-End Controller and Noise-Adding Radiometer, pp. 119-138.

**Buchanan, H. R.**

- 42-101 Long-Range Planning Cost Model for Support of Future Space Missions by the Deep Space Network, pp. 179-190.  
See Sherif, J. S.
- 42-103 Deep Space Network Maintenance Cost Compared to Other Industries and Organizations, pp. 252-261.  
See Remer, D. S.

**Burt, R. W.**

- 42-103 Galileo Array Study Team Report, pp. 161-169.  
See Layland, J. W.

**Carrell, E. A.**

- 42-103 Development of an Improved Analytic Model for the NASA/JPL 70-Meter Antenna Hydraulic Servovalve, pp. 244-251.

R. E. Hill

**Chen, C.-C.**

- 42-104 Steady-State Phase Error for a Phase-Locked Loop Subjected to Periodic Doppler Inputs, pp. 68-73.  
M. Z. Win

**Chen, J. C.**

- 42-104 Analysis of a Thick Dichroic Plate With Rectangular Holes at Arbitrary Angles of Incidence, pp. 9-16.

**Cheung, K.-M.**

- 42-102 A High-Speed Distortionless Predictive Image-Compression Scheme, pp. 73-90.  
P. Smyth and H. Wang
- 42-103 A Node and Frame Synchronization Scheme in the Big Viterbi Decoder Based on Channel Symbol Measurements, pp. 203-214.  
J. I. Statman

- 42-104 Compressed/Reconstructed Test Images for CRAF/Cassini, pp. 88-97.

See Dolinar, S. J.

**Chodas, P. W.**

- 42-103 Optical Self-Crossing Measurements for Ida Ephemeris Improvement, pp. 87-96.  
See Null, G. W.

- Clymer, B. D.**
- 42-103 Active Optics for a Segmented Primary Mirror on a Deep-Space Optical Receiver Antenna (DSORA), pp. 128–134.
- 42-103 Reflected Sunlight Reduction and Characterization for a Deep-Space Optical Receiver Antenna (DSORA), pp. 110–117.
- Cowles, K. A.**
- 42-102 Site Comparison for Optical Visibility Statistics in Southern Arizona, pp. 57–61.
- 42-104 Site Comparison for Optical Visibility Statistics in Southern California, pp.
- Cramer, P. W.**
- 42-104 Initial Studies of Array Feeds for the 70-Meter Antenna at 32 GHz, pp. 50–67.
- Deutsch, L. J.**
- 42-103 A Plan for the Development and Demonstration of Optical Communications for Deep Space, pp. 97–109.
- See Lesh, J. R.
- Dick, G. J.**
- 42-103 Ultrastable Performance of the Superconducting Cavity Maser Oscillator at Short Measuring Times, pp. 21–26.
- See Wang, R. T.
- Djomehri, J.**
- 42-102 Computer Simulations of Ions in Radio Frequency Traps, pp. 10–26.
- See Williams, A.
- Dolinar, S. J.**
- 42-102 The Theoretical Limits of Source and Channel Coding, pp. 62–72.
- F. Pollara
- 42-104 Compressed/Reconstructed Test Images for CRAF/Cassini, pp. 88–97.
- K.-M. Cheung, I. Onyszchuk, F. Pollara, and S. Arnold
- 42-104 Maximum-Entropy Probability Distributions Under  $L_p$ -Norm Constraints, pp. 74–87.
- Edwards, C. D.**
- 42-101 Goldstone Intracomplex Connected Element Interferometry, pp. 1–12.
- Eldred, D. B.**
- 42-101 Application of Inertial Instruments for DSN Antenna Pointing and Tracking, pp. 73–98.
- N. M. Nerheim and K. G. Holmes
- Estefan, J. A.**
- 42-103 Earth-Based Ranging Measurements to a Planetary Radio Beacon: An Approximate Analytic Formulation, pp. 70–86.
- S. W. Thurman and M. P. Batchelder
- Finger, M. H.**
- 42-101 Preliminary Error Budget for an Optical Ranging System: Range, Range Rate, and Differenced Range Observables, pp. 121–135.
- See Folkner, W. M.
- Folkner, W. M.**
- 42-101 Preliminary Error Budget for an Optical Ranging System: Range, Range Rate, and Differenced Range Observables, pp. 121–135.
- M. H. Finger
- 42-102 Navigational Utility of High-Precision Radio Interferometry for Galileo's Approach to Jupiter, pp. 34–46.
- Franco, M. M.**
- 42-103 A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation Part I: Design and Ground Tests, pp. 135–150.
- See Otoshi, T. Y.

- Gawronski, W.**
- 42-104 Predictive Control and Estimation Algorithms for the NASA/JPL 70-Meter Antennas, pp. 30-49.
- Goldstein, R.**
- 42-101 Rings of Earth Detected by Orbital Debris Radar, pp. 191-195.  
L. Randolph
- Gulkis, S.**
- 42-102 Technical Considerations on Using the Large Nançay Radio Telescope for SETI, pp. 152-160.  
F. Biraud, J. Heidmann, and J. Tarter
- Gutierrez-Luaces, B. O.**
- 42-104 Digital Test Signal Generation: An Accurate SNR Calibration Approach for the DSN, pp. 161-174.
- Harabetian, E.**
- 42-102 Computer Simulations of Ions in Radio Frequency Traps, pp. 10-26.  
See Williams, A.
- Heidmann, J.**
- 42-102 Technical Considerations on Using the Large Nançay Radio Telescope for SETI, pp. 152-160.  
See Gulkis, S.
- Hill, R. E.**
- 42-103 Development of an Improved Analytic Model for the NASA/JPL 70-Meter Antenna Hydraulic Servovalve, pp. 244-251.  
See Carrell, E. A.
- Hinedi, S.**
- 42-101 QPSK Loop Lock Detection in the Advanced Receiver, pp. 164-178  
See Mileant, A.
- 42-103 A Comparison of Open-Loop Suppressed-Carrier Acquisition Techniques, pp. 170-188.  
See Shah, B.
- 42-104 The Advanced Receiver II: Telemetry Test Results in CTA 21, pp. 140-156.  
R. Bevan and M. Marina
- 42-104 Overview of Arraying Techniques in the Deep Space Network, pp. 109-139.  
See Mileant, A.
- Hodges, R. E.**
- 42-102 Linear-Phase Approximation in the Triangular Facet Near-Field Physical Optics Computer Program, pp. 47-56.  
See Imbriale, W. A.
- Holmes, J. K.**
- 42-104 A Comparison of Manchester Symbol Tracking Loops for Block V Applications, pp. 175-191.
- 42-104 Open-Loop Frequency Acquisition Suppressed-Carrier Biphasic Signals Using One-Pole Arm Filters, pp. 192-210.  
See Shah, B.
- Holmes, K. G.**
- 42-101 Application of Inertial Instruments for DSN Antenna Pointing and Tracking, pp. 73-98.  
See Eldred, D. H.
- Imbriale, W. A.**
- 42-101 Design and Performance Analysis of the DSS-13 Beam Waveguide Antenna, pp. 99-113  
See Veruttipong, W.
- 42-102 Linear-Phase Approximation in the Triangular Facet Near-Field Physical Optics Computer Program, pp. 47-56.  
R. E. Hodges

- Janik, G. R.**
- 42-101 A Study of the Influence of Hg ( $6^3P_2$ ) Population in a Low-Pressure Discharge on Mercury Ion Emission at 194.2 nm, pp. 61–72.  
See Maleki, L.
- Katow, M. S.**
- 42-101 DSS 15, 45, and 65 34-Meter High-Efficiency Antenna Radio Frequency Performance Enhancement by Tilt Added to the Subreflector During Elevation Angle Changes, pp. 152–163.
- Kent, S. S.**
- 42-102 Interagency Telemetry Arraying for Voyager–Neptune Encounter, pp. 91–118.  
See Brown, D. W.
- Kermode, A. W.**
- 42-104 Design Concepts and Performance of NASA X-Band (7162-MHz/8415-MHz) Transponder for Deep-Space Spacecraft Applications, pp. 247–256.  
See Mysoor, N. R.
- LaBelle, R. C.**
- 42-103 An IF-to-Digital Converter for the Advanced Receiver II (ARX II), pp. 215–243.
- Layland, J. W.**
- 42-103 Galileo Array Study Team Report, pp. 161–169.  
F. D. McLaughlin, P. E. Beyer, D. J. Mudgway, D. W. Brown, R. W. Burt, R. J. Wallace, J. M. Ludwinski, B. D. Madsen, J. C. McKinney, N. A. Renzetti, and J. S. Ulvestad
- Lesh, J. R.**
- 42-103 A Plan for the Development and Demonstration of Optical Communications for Deep Space, pp. 97–109.  
L. J. Deutsch and W. J. Weber
- 42-103 GOPEX: A Deep-Space Optical Communications Demonstration With the Galileo Spacecraft, pp. 262–277.  
See Wilson, K. E.
- Lichten, S. M.**
- 42-104 Demonstration of Precise Estimation of Polar Motion Parameters With the Global Positioning System: Initial Results, pp. 1–8.
- Linfield, R. P.**
- 42-103 Using Planetary Occultations of Radio Sources for Frame-Tie Measurements Part I: Motivation and Search for Events, pp. 1–13.
- Logan, R. T., Jr.**
- 42-102 Design of a Fiber-Optic Transmitter for Microwave Analog Transmission With High Phase Stability, pp. 27–33.  
G. F. Lutes, L. E. Primas, and L. Maleki
- Ludwinski, J. M.**
- 42-103 Galileo Array Study Team Report, pp. 161–169.  
See Layland, J. W.
- Lutes, G. F.**
- 42-102 Design of a Fiber-Optic Transmitter for Microwave Analog Transmission With High Phase Stability, pp. 27–33.  
See Logan, R. T., Jr.
- Madsen, B. D.**
- 42-103 Galileo Array Study Team Report, pp. 161–169.  
See Layland, J. W.
- Maleki, L.**
- 42-101 Inhomogeneous Electric Field Effects in a Linear RF Quadrupole Trap, pp. 51–60.  
See Melbourne, R. K.

42-101 A Study of the Influence of Hg ( $6^3P_2$ ) Population in a Low-Pressure Discharge on Mercury Ion Emission at 194.2 nm, pp. 61–72.

B. J. Blasenheim and G. R. Janik

42-102 Computer Simulations of Ions in Radio Frequency Traps, pp. 10–26.

See Williams, A.

42-102 Design of a Fiber-Optic Transmitter for Microwave Analog Transmission With High Phase Stability, pp. 27–33.

See Logan, R. T., Jr.

#### Malla, R. P.

42-102 The First Geocenter Estimation Results Using GPS Measurements, pp. 1–9.

S. C. Wu

#### Marina, M.

42-104 The Advanced Receiver II: Telemetry Test Results in CTA 21, pp. 140–156.

See Hinedi, S.

#### McKinney, J. C.

42-103 Galileo Array Study Team Report, pp. 161–169.

See Layland, J. W.

#### McLaughlin, F. D.

42-103 Galileo Array Study Team Report, pp. 161–169.

See Layland, J. W.

#### Melbourne, R. K.

42-101 Inhomogeneous Electric Field Effects in a Linear RF Quadrupole Trap, pp. 51–60.

J. D. Prestage and L. Maleki

#### Mellstrom, J.

42-101 Initial Results on Fault Diagnosis of DSN Antenna Control Assemblies Using Pattern Recognition Techniques, pp. 136–151.

See Smyth, P.

#### Mileant, A.

42-101 QPSK Loop Lock Detection in the Advanced Receiver, pp. 164–178

S. Hinedi

42-104 Overview of Arraying Techniques in the Deep Space Network, pp. 109–139.

S. Hinedi

#### Mudgway, D. J.

42-103 Galileo Array Study Team Report, pp. 161–169.

See Layland, J. W.

#### Mysoor, N. R.

42-104 Design Concepts and Performance of NASA X-Band (7162-MHz/8415-MHz) Transponder for Deep-Space Spacecraft Applications, pp. 247–256.

J. D. Perret and A. W. Kermod

#### Nerheim, N. M.

42-101 Application of Inertial Instruments for DSN Antenna Pointing and Tracking, pp. 73–98.

See Eldred, D. H.

#### Nguyen, T. M.

42-102 The Impact of NRZ Data Asymmetry on the Performance of a Space Telemetry System, pp. 139–151.

42-103 A Computational Technique for the Means and Variances of Modulation Losses, pp. 189–202.

#### Null, G. W.

42-103 Optical Self-Crossing Measurements for Ida Ephemeris Improvement, pp. 87–96.

P. W. Chodas

#### Onyszchuk, I.

42-104 Compressed/Reconstructed Test Images for CRAF/Cassini, pp. 88–97.

See Dolinar, S. J.

**Otoshi, T. Y.**

- 42-103 A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation Part I: Design and Ground Tests, pp. 135–150.

S. R. Stewart and M. M. Franco

**Perret, J. D.**

- 42-104 Design Concepts and Performance of NASA X-Band (7162-MHz/8415-MHz) Transponder for Deep-Space Spacecraft Applications, pp. 247–256.

See Mysoor, N. R.

**Pollara, F.**

- 42-102 The Theoretical Limits of Source and Channel Coding, pp. 62–72.

See Dolinar, S. J.

- 42-104 Compressed/Reconstructed Test Images for CRAF/Cassini, pp. 88–97.

See Dolinar, S. J.

- 42-104 Emerging Standards for Still Image Compression: A Software Implementation and Simulation Study, pp. 98–102.

E. S. Arnold

**Prestage, J. D.**

- 42-101 Inhomogeneous Electric Field Effects in a Linear RF Quadrupole Trap, pp. 51–60.

See Melbourne, R. K.

- 42-102 Computer Simulations of Ions in Radio Frequency Traps, pp. 10–26.

See Williams, A.

**Primas, L. E.**

- 42-102 Design of a Fiber-Optic Transmitter for Microwave Analog Signal Transmission With High Phase Stability, pp. 27–33.

See Logan, R. T., Jr.

**Rabkin, J.**

- 42-103 Node and Frame Synchronization in the Big Viterbi Decoder, pp. 154–160.

See Statman, J. L.

**Randolph, L.**

- 42-101 Rings of Earth Detected by Orbital Debris Radar, pp. 191–195.

See Goldstein, R.

**Rauch, L. L.**

- 42-104 On Estimating the Phase of a Periodic Waveform in Additive Gaussian Noise—Part III, pp. 157–160.

**Rayman, M. D.**

- 42-103 GOPEX: A Deep-Space Optical Communications Demonstration With the Galileo Spacecraft, pp. 262–277.

See Wilson, K. E.

- 42-103 Selection and Observability Tests of GOPEX Reference Stars, pp. 278–283.

**Remer, D. S.**

- 42-101 Long-Range Planning Cost Model for Support of Future Space Missions by the Deep Space Network, pp. 179–190.

See Sherif, J. S.

- 42-103 Deep Space Network Maintenance Cost Compared to Other Industries and Organizations, pp. 252–261.

J. S. Sherif and H. R. Buchanan

**Renzetti, N. A.**

- 42-103 Galileo Array Study Team Report, pp. 161–169.

See Layland, J. W.

**Sadr, R.**

- 42-104 Planetary Ephemerides Approximation for Radar Astronomy, pp. 219-234.

M. Shahshahani

**Schwartz, J.**

- 42-103 GOPEX: A Deep-Space Optical Communications Demonstration With the Galileo Spacecraft, pp. 262-277.

See Wilson, K. E.

**Shah, B.**

- 42-103 A Comparison of Open-Loop Suppressed-Carrier Acquisition Techniques, pp. 170-188.

S. Hinedi

- 42-104 Open-Loop Frequency Acquisition Suppressed-Carrier Biphase Signals Using One-Pole Arm Filters, pp. 192-210.

J. K. Holmes

**Shahshahani, M.**

- 42-104 Algebraic Geometric Codes, pp. 103-108.

- 42-104 Planetary Ephemerides Approximation for Radar Astronomy, pp. 219-234.

See Sadr, R.

**Shaik, K.**

- 42-101 A Two-Telescope Receiver Design for Deep Space Optical Communications, pp. 114-120.

- 42-103 Parameter Uncertainties for a 10-Meter Ground-Based Optical Reception Station, pp. 118-127.

**Sherif, J. S.**

- 42-101 Long-Range Planning Cost Model for Support of Future Space Missions by the Deep Space Network, pp. 179-190.

D. S. Remer and H. R. Buchanan

- 42-103 Deep Space Network Maintenance Cost Compared to Other Industries and Organizations, pp. 252-261.

See Remer, D. S.

**Siev, B.**

- 42-103 Node and Frame Synchronization in the Big Viterbi Decoder, pp. 154-160.

See Statman, J. L.

**Smyth, P.**

- 42-101 Initial Results on Fault Diagnosis of DSN Antenna Control Assemblies Using Pattern Recognition Techniques, pp. 136-151.

J. Mellstrom

- 42-102 A High-Speed Distortionless Predictive Image-Compression Scheme, pp. 73-90.

See Cheung, K.-M.

**Solomon, G.**

- 42-103 Generation of Maximum Distance Separable Codes, pp. 151-153.

**Statman, J. I.**

- 42-103 Node and Frame Synchronization in the Big Viterbi Decoder, pp. 154-160.

B. Siev and J. Rabkin

- 42-103 A Node and Frame Synchronization Scheme in the Big Viterbi Decoder Based on Channel Symbol Measurements, pp. 203-214.

See Cheung, K.-M.

**Stewart, S. R.**

- 42-103 A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation Part I: Design and Ground Tests, pp. 135-150.

See Otoshi, T. Y.

**Swanson, L.**

- 42-104 Frame Synchronization for the Galileo Code, pp. 211-218.

See Arnold, E. S.

- Tarter, J.**
- 42-102 Technical Considerations on Using the Large Nançay Radio Telescope for SETI, pp. 152–160.  
See Gulkis, S.
- Thornton, C. L.**
- 42-101 A Reduced-Dynamic Technique for Precise Orbit Determination, pp. 13–25.  
See Wu, S. C.
- Thurman, S. W.**
- 42-101 Applications of Different Design Methodologies in Navigation Systems and Development at JPL, pp. 39–50.
- 42-101 Information Content of a Single Pass of Phase-Delay Data From a Short Baseline Connected Element Interferometer, pp. 26–38.
- 42-103 Deep-Space Navigation With Differenced Data Types Part I: Differenced Range Information Content, pp. 47–60.
- 42-103 Deep-Space Navigation With Differenced Data Types Part II: Differenced Doppler Information Content, pp. 61–69.
- 42-103 DSN Baseline Coordinate and Station Location Errors Induced by Earth Orientation Errors, pp. 40–46.
- 42-103 Earth-Based Ranging Measurements to a Planetary Radio Beacon: An Approximate Analytic Formulation, pp. 70–86.  
See Estefan, J. A.
- 42-103 Navigation and Guidance for a Mars Aerocapture Mission: A System Design Challenge, pp. 27–39.
- Ulvestad, J. S.**
- 42-102 Interagency Telemetry Arraying for Voyager–Neptune Encounter, pp. 91–118.  
See Brown, D. W.
- 42-103 Galileo Array Study Team Report, pp. 161–169.  
See Layland, J. W.
- Veruttipong, W.**
- 42-101 Design and Performance Analysis of the DSS-13 Beam Waveguide Antenna, pp. 99–113  
W. Imbriale and D. Bathker
- Wallace, R. J.**
- 42-103 Galileo Array Study Team Report, pp. 161–169.  
See Layland, J. W.
- Wang, H.**
- 42-102 A High-Speed Distortionless Predictive Image-Compression Scheme, pp. 73–90.  
See Cheung, K.-M.
- Wang, R. T.**
- 42-103 Ultrastable Performance of the Superconducting Cavity Maser Oscillator at Short Measuring Times, pp. 21–26.  
G. J. Dick
- Weber, W. J.**
- 42-103 A Plan for the Development and Demonstration of Optical Communications for Deep Space, pp. 97–109.  
See Lesh, J. R.
- Wee, S.**
- 42-103 GOPEX: A Deep-Space Optical Communications Demonstration With the Galileo Spacecraft, pp. 262–277.  
See Wilson, K. E.
- Williams, A.**
- 42-102 Computer Simulations of Ions in Radio Frequency Traps, pp. 10–26.  
J. D. Prestage, L. Maleki, J. Djomehri, and E. Harabetian

04

- Wilson, K. E.**
- 42-103 GOPEX: A Deep-Space Optical Communications Demonstration With the Galileo Spacecraft, pp. 262-277.  
J. R. Lesh, T.-Y. Yan, J. Schwartz, M. D. Rayman, and S. Wee
- Win, M. Z.**
- 42-104 Steady-State Phase Error for a Phase-Locked Loop Subjected to Periodic Doppler Inputs, pp. 68-73.  
See Chen, C.-C.
- Wu, S. C.**
- 42-101 A Reduced-Dynamic Technique for Precise Orbit Determination, pp. 13-25.  
T. P. Yunck and C. L. Thornton
- 42-102 The First Geocenter Estimation Results Using GPS Measurements, pp. 1-9.  
See Malla, R. P.
- Yan, T.-Y.**
- 42-103 GOPEX: A Deep-Space Optical Communications Demonstration With the Galileo Spacecraft, pp. 262-277.  
See Wilson, K. E.
- Yunck, T. P.**
- 42-101 A Reduced-Dynamic Technique for Precise Orbit Determination, pp. 13-25.  
See Wu, S. C.